

UTILITY APPLICATION

BY

Paul S. Keim

And

Kristen Zinnamon

FOR

UNITED STATES PATENT

ON

DNA FINGERPRINTING FOR CANNABIS SATIVA (MARIJUANA) USING
SHORT TANDEM REPEAT (STR) MARKERS

Docket No.: 130588.91442
EL645096946US

DNA FINGERPRINTING FOR CANNABIS SATIVA (MARIJUANA) USING SHORT TANDEM REPEAT (STR) MARKERS

5

CLAIM TO DOMESTIC PRIORITY

[0001] This application claims benefit of priority to US Provisional application Serial No. 60/397,179, entitled “DNA Fingerprinting For *Cannabis sativa* (Marijuana) Using Short Tandem Repeat (STR) Markers” filed July 19, 2002, by Paul S. Keim et al., and is herein incorporated by reference in its entirety.

10

FIELD OF THE INVENTION

[0002] This invention concerns the molecular analysis of *Cannabis sativa* L. (marijuana) and more specifically provides primer cocktails for multiplex analysis of DNA from purported *Cannabis sativa* L. samples to allow forensic identification and tracking of a leaf sample to its plant source.

15

BACKGROUND

[0003] *Cannabis sativa* L. is one of the oldest crops known to man (Siniscalco Gigliano 2001). Despite its long historical relationship with human civilization, still relatively little is known about the genetic composition of this plant. However, recently many studies have tried to examine the molecular characteristics of *Cannabis* in order to distinguish hemp (fiber) varieties from marijuana (drug) varieties (Gilmore et al. 2003).

20

[0004] The historical and intimate association between *Cannabis sativa* L. (marijuana) and man has no doubt contributed to this plant’s many varieties and uses [1,2]. It is commonly believed that humans introduced *C. sativa* to the Americas in 1545; but before its worldwide introduction, it likely originated and was native to central Asia [3,4]. From even the earliest accounts, man has utilized virtually all parts of the plant for a multitude of purposes, the two most common uses being harvesting the plant for its fiber and drug qualities [5]. The flowers and leaves of the plant are harvested for the chemical resin, delta-9-tetrahydrocannabinol (THC), which when ingested, produces the psychoactive effects that humans experience [6].

25

[0005] A common problem for law enforcement agencies is the correct identification and suppression of illegal growing operations. The forensic community has made significant progress in developing molecular identification techniques for *Cannabis* [7-11]. Virtually all of these experiments have focused on molecular identification methods which exclusively amplify *Cannabis* DNA, enabling forensic investigators to move away from conventional chemical identification tests such as GC-MS, HPLC and histological microscopy. Despite these advances, tests that are capable of individualizing marijuana plants and discriminating between varieties were not available, until recently [12,13]. These kinds of tests are necessary to facilitate the identification and suppression of growing operations by forensic investigators.

[0006] Both Gilmore [12] and Hsieh [13] have investigated the potential utility of short tandem repeat (STR) markers for distinguishing and individualizing *Cannabis* plants. Short tandem repeats (STRs), simple sequence repeats (SSRs), or microsatellites all describe a single type of DNA profiling technology that is useful for providing genetic information about individuals within and among populations. STR genetic markers selectively amplify hypervariable regions of DNA and, when run on gels, generate fluorescent banding patterns that can be used as unique genetic identifiers. Each STR marker is made up of a single DNA sequence, no more than six base pairs long, that is repeated in tandem and individual loci have length polymorphisms in the repeat array [14]. STR markers are useful in forensic investigations because they are polymerase chain reaction (PCR) based and are capable of amplifying small amounts of fairly degraded DNA, which is commonly the condition of biological samples from crime scenes [14]. Additionally, STR markers are desirable because they are a co-dominant marker system and they provide information about the heterozygosity of individual plants.

[0007] Methods and means for reliable and fast genetic analysis of STR markers in *Cannabis sativa* L. have been sought. These analyses would identify purported marijuana samples and would provide a useful forensic tool for linking the source of sample to its plant of origin.

[0008] It is an object of this invention to provide methods and means for STR typing in *Cannabis* to aid forensic investigators in: (i) linking personal possessions of marijuana to plants at the person's residence, (ii) identifying clonally propagated plants as having matching genotypic profiles, and (iii) tracking the distribution patterns of
5 clonally propagated plants within residential areas.

SUMMARY

[0009] The present invention discloses methods and means for detecting and identifying *Cannabis sativa* L. species by short tandem repeat (STR) analysis multiplex genotyping system of STR identified within the genome of *Cannabis sativa* L. STR in the
10 *Cannabis sativa* L. genome are amplified using labeled primers in multiplexed PCRs and electrophoretically separated on polyacrylamide gels for analysis.

[0010] STR loci located throughout the *Cannabis sativa* L. genome have been identified. Isolated nucleic acids having the sequence of STR identified in *Cannabis sativa* L. are presented. In an important aspect of the present invention nucleic acids
15 comprising at least 12, 15, 18 or total consecutive nucleotides of a nucleotide sequence selected from the group consisting of SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 9; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14; SEQ ID NO: 15; SEQ ID NO: 16; SEQ ID NO: 17; SEQ ID NO: 18; SEQ ID NO:
20 19; SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; and sequences complementary thereto are presented.

[0011] In certain preferred embodiments of the invention, these nucleic acids are immobilized on a solid surface and are useful, for example, in the detection of a
25 *Cannabis sativa* L. sample in an assay employing probes, including, but not limited to, a nano-detection device.

[0012] In another important aspect of the invention, primer pairs comprising a forward and a reverse primer are presented for amplification of STR located in DNA from a *Cannabis sativa* L. species. Primer pairs suitable for PCR amplification of STR, by multiplex, may be selected from the group consisting of SEQ ID NO: 1 and 2; SEQ ID

NO: 3 and 4; SEQ ID NO: 5 and 6; SEQ ID NO: 7 and 8; SEQ ID NO: 9 and 10; SEQ ID NO: 11 and 12; SEQ ID NO: 13 and 14; SEQ ID NO: 15 and 16; and SEQ ID NO: 17 and 18; SEQ ID NO: 19 and 20; SEQ ID NO: 21 and 22; SEQ ID NO: 23 and 24; SEQ ID NO: 25 and 26; and SEQ ID NO: 27 and 28.

5 [0013] Combinations of the isolated nucleic acids or primer pairs described herein as “cocktails” are provided for amplification of the STR markers by multiplex. Certain preferred primer pairs have, in addition, an observable group whereby amplified product may be detected. Such groups may be, for example, a fluorescent group or a radioactive group.

10 [0014] In another important aspect of the invention, a method for detecting a *Cannabis sativa* L. species in a sample from a plant, preferably a leaf or flower sample, is presented. The method comprises the steps of:

- i. obtaining DNA from the sample,
- ii. amplifying a STR marker loci in said DNA with a multiplex cocktail
- 15 selected from the group of primer pairs to form amplification products of various sizes and labels; and
- iii. separating amplification products by size and primer label;
- iv. scoring the results of said separation
- v. comparing said scored results to results of analysis of DNA from a
- 20 known species.

[0015] In yet another important aspect of the invention methods for linking a marijuana sample to a plant source are presented. The method comprises the steps of:

- i. determining the identity of DNA in said sample by the present method
- ii. determining the identity of DNA in a sample from a plant by the
- 25 present method; and
- iii. comparing the identities of both samples to determine similarities.

[0016] In another important aspect of the invention, multiplex methods are presented for observing polymorphisms at STR loci in DNA from more than one *Cannabis sativa* L. species to resolve unique genotypes between the species and to allow

30 linking of the sample to its plant of origin. These multiplex methods provide a

convenient and rapid method for genetic discrimination in *Cannabis sativa* L. and, for forensic purposes, provides information necessary to track the source of a purported marijuana sample. Cocktails provided herein are preferably used for amplifying STR in the multiplex methods.

5 [0017] In yet another important aspect of the invention, kits are herein provided for use with commercially available PCR instruments to detect a strain of *Cannabis sativa* L. species. The kits comprise one or more primer pairs suitable for amplifying STR in DNA in a sample of said species by PCR. Preferably the kits comprise primer pairs having SEQ ID NOS: 1-28. Most preferably kits are provided for multiplexing DNA in a
10 sample. These kits comprise primer pair sets, i.e., cocktails, selected from the group of primer pairs.

[0018] The kits may further comprise nucleic acids, enzymes, tag polymerase, for example, salts and buffers suitable for causing amplification by PCR, by multiplex. The kits also comprise preferably a positive control. In certain preferred embodiments of the
15 kit the primers comprise a label whereby amplified STR may be detected. In other preferred embodiments of the kit, labeled nucleic acids are provided. Observable labels are preferably fluorescent molecules or radionucleotides. The kits may also comprise suitable containers and bottles for housing these reagents and or convenient use.

DETAILS

20 [0019] Multiplex methods are presented for rapid genotyping of *Cannabis sativa* L. STR markers described herein provide discriminatory power that enhances the ability of present methods to determine rapidly molecular relationships of *Cannabis sativa* L. samples. A *C. sativa* STR database has been generated by multiplexing 295 samples and eight STR markers. This database illustrates that STR genetic markers in *C. sativa* are
25 both hypervariable and capable of discriminating among individual plants.

[0020] This multiplex typing system is a PCR-based method for genotyping *Cannabis sativa* L. using eight STR loci identified in the present invention. This PCR-based typing system has advantages not present in other PCR-systems: rapid turnaround, amplification with crudely isolated or minute amounts, of DNA. The rapid typing system
30 using eight STR loci has been used to analyze a collection of a 295 samples to detect

genotypic differences between individual *C. sativa* plants. Over 90% of the samples had unique multilocus genotypic profiles and some of the samples with matching profiles were known to be duplicate samples. Although the heterozygosity values detected within this system are fairly low compared to other studies of STRs in plants [12,18], this may
5 be indicative of the selective breeding practices within drug varieties of *C. sativa* plants. It is known that certain drug qualities such as THC content are selectively bred for within this plant [24] and therefore, this system may be detecting some of these highly inbred genotypes. Additional markers, [12,13] would increase the observed heterozygosity values and enhance the power of an STR profiling system for *C. sativa*.

10 [0021] Tri- and tetranucleotide repeat motifs were isolated for their ease of scoring and preferential use in the forensic community [25,26]. Additionally, the observed allele size range (103–364bp) for these markers allows for rapid data collection and accurate scoring due to these smaller fragment sizes [26]. The present system detected 63 alleles. The method of detection may be applied to discover more alleles in
15 other plant samples, including fiber varieties.

[0022] The following definitions are used herein:

[0023] "Polymerase chain reaction" or "PCR" is a technique in which cycles of denaturation, annealing with primer, and extension with DNA polymerase are used to amplify the number of copies of a target DNA sequence by approximately 106 times or
20 more. The polymerase chain reaction process for amplifying nucleic acid is disclosed in US Patent Nos. 4,683,195 and 4,683,202, which are incorporated herein by reference.

[0024] "Primer" is a single-stranded oligonucleotide or DNA fragment which hybridizes with a DNA strand of a locus in such a manner that the 3' terminus of the primer may act as a site of polymerization using a DNA polymerase enzyme.

25 [0025] "Primer pair" is two primers including, primer 1 that hybridizes to a single strand at one end of the DNA sequence to be amplified and primer 2 that hybridizes with the other end on the complementary strand of the DNA sequence to be amplified.

[0026] "Primer site" the area of the target DNA to which a primer hybridizes.

[0027] “Multiplexing” is a capability to perform simultaneous, multiple determinations in a single assay process and a process to implement such a capability in a process is a “multiplexed assay.” Systems containing several loci are called *multiplex* systems described, for example, in US Patent No. 6,479,235 to Schumm, et al., US Patent
5 No. 6,270,973 to Lewis, et al. and 6,449,562 to Chandler, et al.

[0028] “Cocktail” is a mixture of primer pairs selected to amplify one or more STR loci in a multiplex system.

[0029] Isolated nucleic acid” is a nucleic acid which may or may not be identical to that of a naturally occurring nucleic acid. When “isolated nucleic acid” is used to
10 describe a primer, the nucleic acid is not identical to the structure of a naturally occurring nucleic acid spanning at least the length of a gene. The primers herein have been designed to bind to sequences flanking STR loci in *Cannabis sativa* species. It is to be understood that primer sequences containing insertions or deletions in these disclosed sequences that do not impair the binding of the primers to these flanking sequences are
15 also intended to be incorporated into the present invention.

Forensic Utility of STR Markers

[0030] Databases compiled by the present system will be used for drug trafficking and intelligence purposes and to track distribution patterns and growing operations. Additionally, databases are going to be necessary for gaining court acceptance of
20 *Cannabis* DNA fingerprinting systems [12,28].

[0031] Recently, the forensic community has expressed considerable interest in non-human DNA fingerprinting methods for assisting in criminal investigations [27,28]. With the present STR system, forensic investigators will be able to generate genetic profiles of individual *C. sativa* plants and compare them to databases [12,28] or to
25 suspected clonally propagated plants to determine if the profiles match. The identification of clonal growing operations and tracking distribution patterns of individual *Cannabis* plants has the greatest immediate potential for this system. The ability to generate matching genotypic profiles from plants confiscated from independent locations within the same residential area would support the hypothesis that the plants were coming
30 from the same clonal growing operation.

Development of STR Markers

[0032] Of the seven arbitrary repeat motifs that were screened in this protocol, only three (AGC, AAAG, CCT) yielded sequences with sufficient flanking regions for primer development. Over two hundred individual positive clones were sequenced to

- 5 find a total of 33 sequences that contained repeat motifs with at least five repeating units and sufficient flanking sequence on either side of the repeat. Of the 15 markers that were identified as polymorphic, only eight amplified consistently and were easy to score, with minimal stutter problems (Table 2).

Locus Name Dye Label ^a	Repeat Motifs [*]	Aplicon Size Range (bp)	Number of Alleles	Multiplex Mix #
AAAG1 HEX	(AAAG)6	103-135	16	1
ACT1 FAM	(ACT)6	218-224	3	1
AGC8 NED & FAM	(AGC)5	264-279	6	1
AGC9 HEX	(AGC)9	317-335	7	1
AGC1 FAM	(AGC)10	128-164	10	2
AAAG5 NED	(AAAG)5	188-200	4	2
AAAG7 FAM	(AAAG)6	242-266	7	3
AAAG10 FAM	(AAAG)5	352-364	4	3
AGC6 HEX	(AGC)6	200 & 221	2	3
AGC10 NED	(AGC)43	273-327	15	3

- 10 These primer sequences have herein been assigned SEQ ID NO: as follows:

SEQ ID NO

Marker Name

SEQ ID NO: 1

AAAG1

Forward primer

	SEQ ID NO: 2	AAAG1	Reverse primer
	SEQ ID NO: 3	AAAG5	Forward primer
	SEQ ID NO: 4	AAAG5	Reverse primer
	SEQ ID NO: 5	AAAG6	Forward primer
5	SEQ ID NO: 6	AAAG6	Reverse primer
	SEQ ID NO: 7	AAAG7	Forward primer
	SEQ ID NO: 8	AAAG7	Reverse primer
	SEQ ID NO: 9	AAAG10	Forward primer
10	SEQ ID NO: 10	AAAG10	Reverse primer
	SEQ ID NO: 11	AAAG11	Forward primer
	SEQ ID NO: 12	AAAG11	Reverse primer
15	SEQ ID NO: 13	AGC1	Forward primer
	SEQ ID NO: 14	AGC1	Reverse primer
	SEQ ID NO: 15	AGC3	Forward primer
20	SEQ ID NO: 16	AGC3	Reverse primer
	SEQ ID NO: 17	AGC6	Forward primer
25	SEQ ID NO: 18	AGC6	Reverse primer
	SEQ ID NO: 19	AGC8	Forward primer
30	SEQ ID NO: 20	AGC8	Reverse primer
	SEQ ID NO: 21	AGC9	Reverse primer
	SEQ ID NO: 22	AGC9	Reverse primer
35	SEQ ID NO: 23	AGC10	Forward primer
	SEQ ID NO: 24	AGC10	Reverse primer

	SEQ ID NO: 25	ACT1	Forward primer
	SEQ ID NO: 26	ACT1	Reverse primer
5	SEQ ID NO: 27	CCT2	Forward primer
	SEQ ID NO: 28	CCT2	Reverse primer

[0033] The polynucleotides of the present invention may be prepared by two general methods: (1) they may be synthesized from appropriate nucleotide triphosphates, or (2) they may be isolated from biological sources. Both methods utilize protocols well known in the art. The availability of nucleotide sequence information enables preparation of an isolated nucleic acid molecule of the invention by oligonucleotide synthesis. Synthetic oligonucleotides may be prepared by the phosphoramidite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). Complementary segments thus produced may be annealed such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments may be ligated by annealing cohesive termini in the presence of DNA ligase to construct an entire long double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

Total Genetic Diversity

[0034] A total of 295 *C. sativa* samples were analyzed and these samples included representatives from 33 countries or regions around the world. The greatest number of representative samples (188) came from the United States (Table 1). Virtually all of the samples in this study came either from drug confiscations or from known drug varieties of marijuana. Additionally, there were a small number of samples (< 10) that were from known hemp or fiber varieties of *Cannabis*. DNA extracted from four dried samples that came from drug confiscations conducted in 1992 were included in the analyses. Although the DNA was fairly degraded, complete genotypic profiles were obtained for each of these four samples.

[0035] 268 unique genotypes were found from the 295 *C. sativa* samples. For the samples that had at least one matching genotype from a different sample, it was noted that matches corresponded to samples with close geographic locations. All loci amplified robustly using 10 to 15ng DNA and exhibited Mendelian inheritance, with a maximum of two alleles per locus. A total of 63 alleles were detected in this data set, with the number of alleles per locus ranging from two at the AGC6 locus to 16 alleles at the AAAG1 locus (Table 2, Figure 2). The overall observed heterozygosity (averaged across loci) was 0.41 ± 0.01 (mean \pm S.E.) while the expected heterozygosity was calculated to be 0.58 ± 0.05 , when averaged across all eight loci. The average heterozygosity per locus ranged from 0.21 to 0.79.

Allele Frequencies Per Locus

[0036] Figure 2 shows the allele frequencies for each locus in this data set. All observed alleles within each locus, with the exception of two loci, varied by the addition or deletion of single repeat motifs, which is consistent with the assumption that STR loci mutate by insertions and deletions of repeat units. Exceptions of this assumption were observed at the AAAG1 and AGC6 loci. The AAAG1 locus was isolated from a sequence that appeared to contain a 4 bp repeat motif however; samples subjected to the fragment analyses appeared to vary by 2 bp instead of four. The AGC6 locus only had two observable allele sizes, spanning 21bp, which would suggest a mutational event of seven repeat motif units.

[0037] The most diverse marker in this study was the AAAG1 locus, containing 16 alleles and spanning a 32 bp region of the genome, and all expected alleles were observed within this size range (Fig. 2). The second most diverse marker, AGC10 proved to be a noteworthy locus because of its large size range. At this locus we observed 15 alleles and an allelic size range from 273 bp to 336 (Table 2). All but seven of the 22 expected alleles were observed within this 63 bp size range.

Geographic Patterns

[0038] A neighbor-joining tree based on the proportion of shared alleles between samples was constructed. An assignment test was conducted to explore the potential utility of these markers for making geographic assignments based on a particular

genotype. The results suggest a possible utility of these markers in detecting geographic differences on large, regional scales such as continents. The results of the neighbor-joining tree (Fig. 3) depict large-scale geographic clustering based on similar genotypes. All states within North America clustered together. Additionally, samples from Europe
5 and Asia clustered together, while samples from South America and Africa clustered together.

[0039] The results of the assignment test (Fig. 4) indicate that in general, genotypes can be correctly assigned to the right continent at least 50% of the time. Genotypes from the African population (13 samples) were correctly assigned to Africa in
10 all instances; whereas genotypes from the Asian population (46 samples) were only correctly assigned to Asia 61% of the time (Table 1, Fig. 4). The North American population had the largest sample size (196 samples) and their genotypes were correctly assigned 72% of the time. This North American population, with its relatively large sample size, suggests that correct assignments to populations may increase with
15 increasing sample size.

Genetic Diversity Among Individual Samples

[0040] We conducted an analysis of molecular variance (AMOVA) to determine the distribution of the genetic variation. Our findings revealed that the greatest proportion of genetic variation (~ 90%) was among individual samples, within counties
20 and states (Table 3). While the AMOVA did indicate that there were significant differences ($P < 0.0001$) within countries and continents, this variation only accounted for approximately 8% of the total variance. This analysis also shows that the variation among the continents was not statistically significant at 2% (Table 3). The results of the AMOVA (Table 3) suggest that these markers are able to detect genetic differences
25 between individual samples. Additionally, the number of unique genotypes observed, 268 out of 295 samples, also indicates that this system is capable of detecting a sizeable portion of the variation in the samples analyzed.

EXPERIMENTAL DETAILS

DNA Extraction and Sample Preparation

[0041] *Cannabis sativa* DNA was extracted from dried leaf and flower material, in crime laboratories independent of our laboratory, by criminalistics professionals licensed to legally handle these plant samples. Virtually all of the samples came from drug confiscations or from known drug varieties of marijuana. Four different crime laboratories provided DNA samples for this study and there were two main extraction protocols that these agencies used. From these laboratories, we obtained a total of 295 samples with a wide geographic distribution, including representative samples from five different continents (see Table 1). For samples within the United States, the sample location generally refers to the location of the drug confiscation and cultivation. However, the international sample locations do not necessarily correspond to the location of cultivation. Rather these locations correspond to region where the seeds were obtained.

15 [0042] The majority of samples (240 samples) were extracted by the Appalachian H.I.D.T.A. Marijuana Signature Laboratory, Frankfort, KY, using a modified CTAB (cetyltrimethylammonium bromide) protocol described by Weising *et al.* [15]. The remaining 55 samples were extracted in three independent laboratories, all using QIAGEN®'s DNeasy® plant mini kit (QIAGEN, Inc., Valencia, CA, USA), following
20 manufacturers recommendations for dried plant material. DNA samples were received in 100-150 µl of TE buffer [10 mM tris-HCl at pH 8.0, 1 mM EDTA (ethylenediaminetetraacetic acid)] and stored at –20 °C. The approximate yield of each sample was assessed on a 0.7% agarose gel, where samples were compared to a Lambda Hind III DNA mass ladder of known concentrations (Invitrogen, Carlsbad, CA, USA).
25 All DNA samples were then diluted to approximately 10 to 15 ng/ul for the subsequent analyses.

Development of STR Markers

[0043] The STR (microsatellite) markers were developed using a modified magnetic bead protocol that was first described by Li *et al.* [16] and modified by Pearson
30 [17]. Genomic DNA was digested from three different marijuana plants using an *MboI* restriction enzyme (Invitrogen; Carlsbad, CA). Sau 3a I Linkers A and B (SAULA: 5'

GCG GTA CCC GGG AAG CTT GG 3' and SAULB: 5' GAT CCC AAG CTT CCC GGG TAC CGC 3') were ligated onto the digested genomic DNA and SAULA was used as a primer for subsequent polymerase chain reactions (PCR) [16]. The digested genomic DNA was amplified in multiple PCR reactions and concentrated to gain enough DNA for the following bead hybridization process.

[0044] Seven arbitrary repeat motifs were chosen as probes for the bead hybridization reactions based on a review by Cardle et al. [18] where they suggested that plants contain more AT-rich repeats than GC-rich repeats. The short tandem repeat (STR) probes were ordered from Integrated DNA Technologies (Coralville, IA, USA) with a biotin label on the 5' end of the probes [(AGC)₈, (AAAG)₅, (CCT)₈, (AATT)₅, (ATT)₈, (GATA)₅, (ATGC)₅]. These repeat probes were then added to a bead hybridization reaction to select for fragments of DNA that contain the repeat motif of the probe. The goal of this bead hybridization process was to allow the fragments containing repeats to anneal to the biotin-labeled probes. After the hybridization, the selected fragments were isolated from the rest of the genomic DNA using streptavidin coated magnetic beads, which bind to the biotin labeled probes. These fragments were then eluted and re-amplified using the SAULA primer in additional PCR reactions. The bead hybridization and PCR re-amplification processes were then repeated two additional times to enrich for genomic DNA containing the selected repeats.

[0045] Once the bead hybridization and selection process was completed, the repeat enriched DNA was then ligated into a pGEM-T vector from ProMega (Madison, WI, USA) in order to begin the sequencing phase of this protocol. The vectors were cloned into electrocompetent *E. coli* cells that were then plated onto selective media containing [0.1 mg/mL ampicillin, 0.05 mg/mL X-Gal, and 1mM IPTG] and positive clones were sequenced on an ABI PRISM[®] 377 DNA Sequencer (Applied Biosystems; Foster City, CA, USA). The sequencing reactions were standard 20 µl reactions using the ABI PRISM[®] BigDye[™] Terminators sequencing kits (Applied Biosystems; Foster City, CA, USA) and 3.2 pmol of PCR product for template. Sequences containing repeat motifs and sufficient flanking sequence were used to design primers with PrimerSelect software (DNASTAR Inc.; Madison, WI, USA).

[0046] Thirty-three primer pairs were screened on 3% agarose gels against 24 samples from different locations to identify polymorphic markers. Of the 33 markers that were initially screened, fifteen were determined to be polymorphic and we obtained these 15 markers with fluorescent dye labels. The fluorescent markers were tested on an ABI PRISM® 377 DNA Sequencer (Applied Biosystems; Foster City, CA, USA) and seven of the 15 markers were eliminated due to problems with scoring or very low levels of polymorphism. The remaining eight markers (see Table 2) were tested in three multiplex reactions with two to four markers per mix and gels were run using GeneScan 2.1.1 (Applied Biosystems; Foster City, CA, USA) collection software on an ABI PRISM® 377 DNA Sequencer (Applied Biosystems; Foster City, CA, USA). Once multiplex reactions were optimized, 295 samples from individual plants were screened across all eight markers.

PCR Amplification and Fragment Analysis

[0047] The eight STR markers were optimized to amplify DNA in three 10 µl multiplex reactions (see Table 2). The multiplex mixes each contained approximately 10-15 ng of template from *C. sativa* in a 10 µl PCR including the following (final concentrations): 1X PCR buffer (Invitrogen; Carlsbad, CA, USA), 3 mM MgCl₂ (Invitrogen; Carlsbad, CA, USA), 200 µM dNTPs, 0.2 µM fluorescent forward primers, 0.2 µM unlabeled forward primers, 0.4 µM unlabeled reverse primers, and 1 unit Platinum DNA Taq Polymerase (Invitrogen; Carlsbad, CA, USA). Amplification reactions were then carried out in 96-well microplates in a DNA engine thermocycler (MJ Research, Inc.; Waltham, MA, USA) and the reaction contained a total of 35 cycles. The thermocycling conditions were as follows: an initial incubation of 95°C for 5 min, next a cycle of denaturing at 95°C for 30 sec, annealing at (59°C, 60°C, or 62°C) for 30 sec, and extending at 72°C for 30 sec, repeated for a total of 35 cycles, with a final extension of 72°C for 2 min, and ending with a holding temperature of 15°C.

[0048] The PCR products were then diluted 1:10 with E-pure® purified water in preparation for fragment analysis on the ABI PRISM® 377 DNA Sequencer (Applied Biosystems; Foster City, CA, USA). A size standard ladder mix was prepared with 0.75 µl deionized formamide, 0.25 µl of ROX labeled MapMarkers™1000 (BioVentures, Inc.;

Murfreesboro, TN, USA), and 0.1 µl of blue dextran loading dye (supplied with the ROX size ladder). Approximately 1 µl of the size standard ladder mix was added to 1 µl of the diluted amplification products and denatured at 95°C for 2 minutes. From this mixture, roughly 1.6 µl was loaded on a porous membrane comb (The Gel Company; San Francisco, CA, USA) and then electrophoresed in a 5% polyacrylamide gel on the ABI PRISM® 377 DNA Sequencer (Applied Biosystems; Foster City, CA, USA) for 3.5 hours.

Scoring of STR Loci and Data Analysis

[0049] Electrophoresis data was collected automatically with GeneScan™ 2.1.1 software (PE Applied Biosystems; Foster City, CA, USA); following collection, this software was also used to determine the allele sizes by implementing the local Southern method.

[0050] After initial scoring was completed, Genotyper™ software (Applied Biosystems; Foster City, CA, USA) was used to confirm the allele scores. Banding patterns of homozygous and heterozygous genotypes were consistent with that of a single peak for homozygotes and double peaks for heterozygotes. Once all of the data scoring was complete, random samples were re-amplified and independently re-run to assess reproducibility and confirm the scoring and banding patterns.

[0051] Statistical analyses of the data were performed using a multitude of different analysis packages. An Excel add-in called The Excel Microsatellite Toolkit V3.1 [19] was used to calculate the number of matching genotypes, number of alleles, allele frequencies, and observed and expected heterozygosity. A distance matrix was generated in MICROSAT [20] based on the proportion of shared alleles, which was then input into PHYLIP [21] to construct a phylogenetic tree using a neighbor-joining algorithm. Genetic differentiation among continents was calculated in Arlequin V2.0 [22] using an Analysis of Molecular Variance (AMOVA). Finally an assignment test was performed in GenAlEx V5 [23].

EXAMPLES

The following examples illustrate locus sequences for all fifteen polymorphic loci isolated from *Cannabis sativa*. **Forward** and **Reverse** primers are underlined.

Variable regions are in **lower case**. *Most probes have an additional G added to the 5' end of the oligo to increase adenylation. All sequences are 5' → 3'

EXAMPLE 1

This example illustrates the amplicons produced during the amplification of STR locus **AAAG 1** with multiplex cocktails comprising primer pairs SEQ ID NO: 1 and SEQ ID NO: 2.

Sequence for **AAAG 1** locus:

GCGGTACCCGGAAGCTTGGGATCTAAACTGAGAGGTGGGTTTTGGTCAGAA
AGCGAAGACCTTTAGACCCAATATGAAGGAGaagaagaagaagaagaagaagaa
gaaagaagaagaagaagaagAAAACACAGCTAGCAAAAGAAGTAAAGACAGGCAG
15 CCATCATTAATGGCAGAGAGATAGAGTGAGAAAGAGATAGAAAGGAGGAG
AGAGAGAGAGATAGAGAGTACAAGAAAGAAAGAGCAAAGCCAAGCTTCCCCG
GGTACCGC

AAAG1F: GTCAGAAAGC GAAGACCTTT AGA [23bp]

20 AAAG1R: GTAAAGACAG GCAGCCATC [19bp]

AAAG1F (rev. comp.): TCTAAAGGTC TTCGCTTTCT GAC [23bp]

AAAG1R (rev. comp.): GATGGCTGCC TGTCTTTAC [19bp]

25 AAAG1 array: AAGAAGAAGA AGAAGAAGAA GAAGAAAGAA AGAAAGAAAG
AAAGAAAG [48bp]

AAAG1 motif: (AAG)8 + (AAAG)6

AAAG1 amplicon: [275bp]

GCGGTACCCG GGAAGCTTGG GATCTAACT GAGAGGTGGG
TTTTGGTCAG AAAGCGAAGA CCTTTAGACC CAATATGAAG
GAGAAGAAGA AGAAGAAGAA GAAGAAGAAA GAAAGAAAGA
AAGAAAGAAA GAAAACACAG CTAGCAAAAG AAGTAAAGAC
5 AGGCAGCCAT CATTAATGGC AGAGAGATAG AGTGAGAAAAG
AGATAGAAAAG GAGGAGAGAG AGAGAGATAG AGAGTACAAG
AAAGAAAGAG CAAAGCCAAG CTTCCCGGGT ACCGC

AAAG1 (reverse compliment): [275bp]

10 GCGGTACCCG GGAAGCTTGG CTTTGCTCTT TCTTTCTTGT ACTCTCTATC
TCTCTCTCTC TCCTCCTTTC TATCTCTTTC TCACTCTATC TCTCTGCCAT
TAATGATGGC TGCCTGTCTT TACTTCTTTT GCTAGCTGTG TTTTCTTTCT
TTCTTTCTTT CTTTCTTTCT TCTTCTTCTT CTCTTCTTC TTCTCCTTCA
TATTGGGTCT AAAGGTCTTC GCTTTCTGAC CAAAACCCAC CTCTCAGTTT
15 AGATCCCAAG CTTCCCGGGT ACCGC

EXAMPLE 2

This example illustrates the amplicons produced during the amplification of STR
locus **AAAG 5** with multiplex cocktails comprising primer pairs SEQ ID NO: 3 and SEQ
20 ID NO: 4.

Sequence for **AAAG 5** locus:

GCGGTACCCGGGAAGCTTGGCATCAACTTGTCAAGCATTTAATATAAGATTG
GAATATATGTAACATCTTCAATTAATGCTTATAGCCCATATGTTTTCTACTA
CTTCTTCTTTTTTCAGTTGGTGTATATAGCTTGATGATTACTTTCACGGGTGTaaa
25 caaagaagaagaagaagaagaagaagaagaagaagACATGGGTGAGCTGCTTCTGTATATG
TTGTTCCATGGAGAACAAGAAGAAACAAAGTATTCCTGAAGTTGTGATAT
TTGTACCTTCATTGAAAATACCATTACAATCTGATCCCAAGCTTCCCGGGTAC
CGC

30 AAAG5F: TCAATTAATG CTTATAGCCC ATATGTTTTT TACTAC [36bp]
AAAG5R: AGAACAAGAA GAAACAAAGT ATTCCTGAAG TTG [33bp]

AAAG5F (rev. comp.): GTAGTAGAAA ACATATGGGC TATAAGCATT
AATTGA [36bp]

5 AAAG5R (rev. comp.): CAACTTCAGG AATACTTTGT TTCTTCTTGT TCT
[33bp]

AAAG5 array: AAACAAAAGA AGAAGAAAGA AAGAAAGAAA GAAAGAAG
[48bp]

10 AAAG5 motif: (AAAC)1 + (AAAAG)1 + (AAG)2 + (AAAG)5 + (AAG)1

AAAG5 amplicon: [327bp]

GCGGTACCCG GGAAGCTTGG CATCAACTTG TCAAGCATTT
AATATAAGAT TGGAATATAT GTAACATCTC AATTAATGCT TATAGCCCAT
ATGTTTTCTA CTA CTCTTCTTC TTTTCAGTT GGTGTTATAT AGCTTGATGA
15 TTACTTTTAC GGTGTAAACA AAAGAAGAAG AAAGAAAGAA
AGAAAGAAAG AAGACATGGG TTGAGCTGCT TCTGTATATG
TTGTTCCATG GAAGAACAAG AAGAAACAAA GTATTCCTGA
AGTTGTGATA TTTGTACCTT CATTGAAAAT ACCATTACAA TCTGATCCCA
AGCTTCCCGG GTACCGC

20 AAAG5 reverse compliment: [327bp]

GCGGTACCCG GGAAGCTTGG GATCAGATTG TAATGGTATT
TTCAATGAAG GTACAAATAT CACAACTTCA GGAATACTTT GTTTCTTCTT
GTTCTTCCAT GGAACAACAT ATACAGAAGC AGCTCAACCC ATGTCTTCTT
25 TCTTTCTTTC TTTCTTTCTT CTTCTTTTGT TTACACCGTG AAAGTAATCA
TCAAGCTATA TAACACCAAC TGAAAAAGAA GAAGTAGTAG
AAAACATATG GGCTATAAGC ATTAATTGAG ATGTTACATA TATTCCAATC
TTATATTAAA TGCTTGACAA GTTGATGCCA AGCTTCCCGG GTACCGC

30 **EXAMPLE 3**

This example illustrates the amplicons produced during the amplification of STR locus **AAAG 6** with multiplex cocktails comprising primer pairs SEQ ID NO: 5 and SEQ ID NO: 6.

Sequence for **AAAG 6** locus:

5 GCGGTACCCGGGAAGCTTGGCTTAGATTAAGAATATTTGTAGTTTCGTA
TATTCTTGCCTTTTTCAAGATTTCTT
GCTTGTTTAGGGTATCTGCCATTTTCTTCTCCTTTCAGAGCTTCTTCTAATC
CAAGATTCCCAAGATGAGCAATTGTC
TTTTCACCCACAGACTGAAATTGTT**TTTGCCATTGATTTCTCCTCCTCAT**
10 **A**CTTCTCCAAAGACATTATTGAACAAATAAG**Gaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag**
aagaaagAAAACTTATGGCCAGTAAGCGTTTCCCTTGTTGGTTACCTTTCTTCA
GTCTTTGAGGAATTCATTCGAACACTCTGTCAACCTCAACTGGTTTCTTCAAA
CTCTAATCTGAAACCTGGCTCTTGATAACAGTTTGTGAGGATTGGTCTCCTCT
TCTCCAATCTC**AGATCCCAAGCTTCCCGGGTACCGC**

15

AAAG6F: TTTGCCATTG ATTCCTCCT CTCATAC [28bp]

AAAG6R: AGATCCCAAG CTTCCCGGGT ACC [23bp]

AAAG6F (rev. comp.): GTATGAGGAG GAGGAAATCA ATGGCAAA [28bp]

AAAG6R (rev. comp.): GGTACCCGGG AAGCTTGGA TCT [23bp]

20 AAAG6 array: AAAGAAAGAA AGAAAGAAAG AAAGAAAGAA AGAAAG
[36bp]

AAAG6 motif: (AAAG)⁹

AAAG6 locus: [469bp]

25 GCGGTACCCG GGAAGCTTGG CTTAGATTAA GAATATTTGT AGTTTCGTAC
TTGTATTCCT TGCCTTTTTC AAGATTTCTT GCTTGTTTAG GGTATCTGCC
ATTTTCTTT CTCCTTTCAG AGCTTCTTCT AATCCAAGAT TCCCAAGATG
AGCAATTGTC TTTTCACCCC ACAGACTGAA ATTGTTTTTG CCATTGATTT
CCTCCTCCTC AACTTCTCC AAAGACATTA TTGAACAAAT
30 AAGAAAGAAA GAAAGAAAGA AAGAAAGAAA GAAAGAAAGA
AAAACTTATG GCCAGTAAGC GTTTCCTTG TTGGTTACCT TTCTTCAGTC

AAAG6 reverse compliment: [469bp]

15 TACGAAACTA CAAATATTCT TAATCTAAGC CAAGCTTCCC GGGTACCGC

EXAMPLE 4

20 ID NO: 8.

Sequence for **AAAG 7** locus:

25 **gaaagaaag**ATAGATAGATACCTGGTAGTGGGTTGGTTGGTTGGTGGTGATGAGT

AAAG7F: CTACAAAGAT TCCCACACTC AATAATGCAA ATACAA [36bp]

AAAG7R: AGTAAGGATT TGGTTTTTCGG CTTTCGTTCT T [31bp]

AAAG7F (rev. comp.): TTGTATTTGC ATTATTGAGT GTGGGAATCT TTGTAG
[36bp]

5 AAAG7R (rev. comp.): AAGAACGAAA GCCGAAAACC AAATCCTTAC T[31bp]

AAAG7 array: AAAACAAAAA GAAAAGAAAG AAAGAAAGAA AG [32bp]

AAAG7 motif: (AAAAAG)1 + (AAAAG)1 + (AAAG)4

AAAG7 locus: [434bp]

10 GCGGTACCCG GGAAGCTTGG ATCAGAAAGA CAAGACAAGA

TAGGGACTAC TACAAAGATT CCCACACTCA ATAATGCAAA

TACAATTATT AGTACTAATA ATGAAAACAA CATCAAATTA

AAGAAAAACC ATAGAAGAAA AAAAAAGAA AAGAAAGAAA

GAAAGAAAGA TAGATAGATA CCTGGTAGTG GGTTGGTTGG

15 TTGGTGGTGA TGAGTACTGA AATGGAAGAC AATGAAAGGA

GAAGGGGTTT ACAGTGTTAA CACTATAGTA AGGATTTGGT TTTCGGCTTT

CGTTCTTTTA AGGAAGATGG GTGTTTGAGA ATGGATTGAG

TAGTACAAGT CCAAATTCAC AAGCAATTGC AGAGGCAGAC

GATGACTTCT TCAAATTCAT AAGCAAGTGC

20 CGAGGCAACC GATCCCAAGC TTCCCGGGTA CCGC

AAAG7 reverse compliment: [434bp]

GCGGTACCCG GGAAGCTTGG GATCGGTTGC CTCGGCACTT

GCTTATGAAT TTGAAGAAGT CATCGTCTGC CTCTGCAATT GCTTGTGAAT

TTGGACTTGT ACTACTCAAT CCATTCTCAA ACACCCATCT TCCTTAAAAG

25 AACGAAAGCC GAAAACCAAA TCCTTACTAT AGTGTTAACA

CTGTAAACCC CTTCTCCTTT CATTGTCTTC CATTTCAGTA CTCATCACCA

CCAACCAACC AACCCACTAC CAGGTATCTA TCTATCTTTC TTTCTTTCTT

TCTTTTCTTT TTGTTTTCTT CTATGGTTTT TCTTTAATTT GATGTTGTTT

TCATTATTAG TACTAATAAT TGTATTTGCA TTATTGAGTG TGGAATCTT

30 TGTAGTAGTC CCTATCTTGT CTTGTCTTTC TGATCCAAGC TTCCCGGGTA

CCGC

EXAMPLE 5

This example illustrates the amplicons produced during the amplification of STR locus **AAAG 10** with multiplex cocktails comprising primer pairs SEQ ID NO: 9 and SEQ ID NO: 10.

5 Sequence for **AAAG 10** locus:

GCGGTACCCGGAAGCTTGGATAA**CAAAAATTCATACATAAGGCACGAAG**
AGATAGACATAGGaaagaaagaaagaaagaaagGAAAAAAAAAATACTAAAACGAC
ATACACGGTCTTAGAGGACGAAGCAACTGCGCCGCCGCCGGTGACTGGGTTC
CT
10 TGGTCGAGAGGGAAAAAGAGGTTTTTGGTCTCTCTGACTCTGTTGTGCAGTGA
GATGAGGAGTGGAGAGTCGGATAGCATCATTTTTTACACTAACTGAGAAGAAC
AACTTTTGATTTGGTTTGGTTTAAGGAAGAAAAAATCCACATCGACTTGTTA
TAGCTTTTTTAATATGTTTATATTGATTAC**TTTATACAGTCCTATCGCCGGG**
TCCAAGGCTTCCCGGGTACCGC

15

AAAG10F: CAAAATTCA TACATAAGGC ACGAAGAGAT AGACA [35bp]
AAAG10R: TTTATACAGT CCTATCGCCG GGTCCAA [27bp]

20

AAAG10F (rev. comp.): TGTCTATCTC TTCGTGCCTT ATGTATGAAT TTTTG
[35bp]
AAAG10R (rev. comp.): TTGGACCCGG CGATAGGACT GTATAAA [27bp]

25

AAAG10 array: AAAGAAAGAA AGAAAGAAAG [20bp]
AAAG10 motif: (AAAG)⁵

30

AAAG10 locus: [391bp]
GCGGTACCCG GGAAGCTTGG ATAACAAAAA TTCATACATA
AGGCACGAAG AGATAGACAT AGAAAGAAAG AAAGAAAGAA
AGGAAAAAAAA AAAATACTAA AACGACATAC ACGGTCTTAG
AGGACGAAGC AACTGCGCCG CCGCCGGTGA CTGGGTTCCT
TGGTCGAGAG GGAAAAAGAG GTTTTTGGTC TCTCTGACTC TGTTGTGCAG

TGAGATGAGG AGTGGAGAGT CGGATAGCAT CATTTTACAC
CTAACTGAGA AGAACAACTT TTGATTGTTT TTGGTTTAAG
GAAGAAAAAA TCCCACATCG ACTTGTTATA GCTTTTTTAA TATGTTTATA
TTGATTACTT TATACAGTCC TATCGCCGGG TCCAAGCTTC CCGGGTACCG
5 C

AAAG10 reverse compliment: [391bp]
GCGGTACCCG GGAAGCTTGG ACCCGGCGAT AGGACTGTAT
AAAGTAATCA ATATAACAT ATAAAAAAG CTATAACAAG
10 TCGATGTGGG ATTTTTTCTT CCTTAAACCA AACCAAATCA AAAGTTGTTC
TTCTCAGTTA GTGTAAAAAT GATGCTATCC GACTCTCCAC TCCTCATCTC
ACTGCACAAC AGAGTCAGAG AGACCAAAAA CCTCTTTTTC
CCTCTCGACC AAGGAACCCA GTCACCGGCG GCGGCGCAGT
TGCTTCGTCC TCTAAGACCG TGTATGTCGT TTAGTATTT TTTTTTTCC
15 TTTCTTTCTT TCTTTCTTC TATGTCTATC TCTTCGTGCC TTATGTATGA
ATTTTTGTTA TCCAAGCTTC CCGGGTACCG C

EXAMPLE 6

This example illustrates the amplicons produced during the amplification of STR locus
20 **AAAG 11** with multiplex cocktails comprising primer pairs SEQ ID NO: 11 and SEQ ID
NO: 12.

Sequence for **AAAG 11** locus:

TTGCGGTACCCGGGAAGCTTGGATCTTAAAAGTTCAGGGGGGCAAAAATCATA
ATTAGCCTATTGTTAATAATAGACCCTCCTAAAAATCGTTTTGCAAAATAACA
25 TTCTTTTCATAATTGTTTGCAAAATAATCTTTCTCTAGAATCCAAATAGTAT
TGAGAATTTTAAACAAAGTATTTGGAATTCTTAACAAAATGTTAGATTGTGAA
GGTGCTAGAAAGGTCATTTTTTGTAAAAATTATCATCTATCAATTACTCATG
ATAGATTGTTGGAATAGAATCACAAGTTTTTGTACACTATTATGTGGAGTGA
TTGGTGAAAATACACTTATTATGCAAATTGTACATAAAAAGAAGGaaagaaagaa
30 **agaaag**TCTATTTACCAAACAAAAGAAACACCTTTATTATGTGAAAGTGATTG
ATGCATAAAGACTAATAATGCAGGATTTGAAGAGCCTTTGAGAGCAT**GTTGT**

GGTCATGGTGGGAAGTATAATTTTAATAAGAaCATTGGATGTGGGGGCAAG
AAAATGGTCCATGGGAAGAGATTTTGGTGGGAAAGGCTTGTAAGATCCAA
GCTTCCCGGGTACCGC

- 5 AAAG11F: TTTTCATAAT TGTTTGCAAA ATAATCTTTC TCTAGAA [37bp]
AAAG11R: GTTGTGGTCA TGGTGGGAAG TATAATTTTA ATA [33bp]

AAAG11F (rev. comp.): TTCTAGAGAA AGATTATTTT GCAAACAATT
ATGAAAA [37bp]

- 10 AAAG11R (rev. comp.): TATTAAAATT ATACTTCCCA CCATGACCAC AAC
[33bp]

AAAG11 array: AAAGAAAGAA AGAAAG [16bp]
AAAG11 motif: (AAAG)₄

15

AAAG11 locus: [596bp]

- TTGCGGTACC CGGGAAGCTT GGATCTTAAA AGTTCAGGGG
GCAAAAATCA TAATTAGCCT ATTGTTAATA ATAGACCCTC CTAAAAATCG
TTTTGCAAAA TAACATTCTT TTCATAATTG TTTGCAAAAT AATCTTTCTC
20 TAGAATCCAA ATAGTATTGA GAATTTTAA CAAAGTATTT GGAATTCTTA
ACAAAATGTT AGATTGTGAA GGTGCTAGAA AGGTCATTTT
TTGTAAAAA TTATCATCTA TCAATTACTC ATGATAGATT GTTGAATAG
AATCACAAGT TTTTGTTACA CTATTATGTG GAGTGATTGG TGAAAATACA
CTTATTATGC AAATTGTACA TAAAAAGAAG GAAAGAAAGA
25 AAGAAAGTCT ATTCACCAA ACAAAGAAA CACCTTTATT
ATGTGAAAGT GATTGATGCA TAAAGACTAA TAATGCAGGA
TTTGAAGAGC CTTTGAGAGC ATGTTGTGGT CATGGTGGGA AGTATAATTT
TAATAAGAAC ATTGGATGTG GGGGCAAGAA AATGGTCCAT
GGGAAAGAGA TTTTGGTGGG
30 AAAGGCTTGT AAAGATCCAA GCTTCCCGGG TACCGC

AAAG11 reverse compliment: [596bp]

GCGGTACCCG GGAAGCTTGG ATCTTTACAA GCCTTTCCCA CCAAATCTC
TTTCCCATGG ACCATTTTCT TGCCCCCACA TCCAATGTTC TTATTAAAAT
TATACTTCCC ACCATGACCA CAACATGCTC TCAAAGGCTC TTCAAATCCT
GCATTATTAG TCTTTATGCA TCAATCACTT TCACATAATA AAGGTGTTTC
5 TTTTGTTTGG TGAAATAGAC TTTCTTTCTT TCTTTCCTTC TTTTATGTA
CAATTTGCAT AATAAGTGTA TTTTCACCAA TCACTCCACA TAATAGTGTA
ACAAAACTT GTGATTCTAT TCCAACAATC TATCATGAGT AATTGATAGA
TGATAATTTT TAACAAAAAA TGACCTTTCT AGCACCTTCA CAATCTAACA
TTTTGTAAAG AATTCCAAAT ACTTTGTAA AAATTCTCAA TACTATTTGG
10 ATTCTAGAGA AAGATTATT TGCAAACAAT TATGAAAAGA ATGTTATTTT
GCAAACGAT TTTTAGGAGG GTCTATTATT AACAATAGGC TAATTATGAT
TTTTGCCCC TGAACTTTAA AGATCCAAGC TTCCCGGGTA CCGCAA

EXAMPLE 7

15 This example illustrates the amplicons produced during the amplification of STR
ocus **AGC 1** with multiplex cocktails comprising primer pairs SEQ ID NO: 13 and SEQ
ID NO: 14.

Sequence for **AGC 1** locus:

GGGCCCCGACGTCGCATGCTCCCGGCCGCGCCATGGCCGCGGGATTACCCGGGA
20 AGCTTGGATAAGACCATGGCAAGAAAAGATGAGCAACAGAATGTGGTAATT
CAATACAAACAGAACACAAGTCGAATGGATAATAATAAGAAGAAACAG
TTGCCAAGCTGTCAAAGAAATCACAGAACAATTTAGAGTTACAACAACCAT
TCGTGCCTGGAAAATTAGTATCACAAGATAATGGAAAACAAGTTTTACAGAC
AAGAAAACAAAAGGGTAGCACTGGTAGTAGTGAAGTTATGG**CAAAGAGTGT**
25 **ATCGAAACCTGT**CCGTGATGGAACAAATTTCAACAGAA**gcagcagcagcagcagca**
gcagcagcagcCACAGTCTAACCAAGAAAAGTTGAATAAGAAAGGTTTGAAAAAA
GGTACTAATACAGACGATGTGGTGGGGTAGAAAGAAATTTGGCTGAATC
CAATTCGTTAAGGAATACAACAATCGAAGCCCGGATCCCAAGCTTCCCGGG
TACCGC

30

AGC1F: CAAAGAGTGT ATCGAAACCT GTC [23bp]

AGC1R: GTACTAATAC AGACGATGTG GTGGG [25bp]

AGC1F (rev. comp.): GACAGGTTTC GATACACTCT TTG [23bp]

AGC1R (rev. comp.): CCCACCACAT CGTCTGTATT AGTAC [25bp]

5

AGC1 array: AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC [30bp]

AGC1 motif: (AGC)₁₀

AGC1 locus: [529bp]

10 GGGCCCGACG TCGCATGCTC CCGGCCGCCA TGGCCGCGGG
ATTACCCGG GAAGCTTGGA TAAGACCATG GCAAGAAAAG
ATGAGCAACA GAATGTGGTA ATTCAATACA AACAGAACAC
AAGTCGAATG GATAATAATA ATAAGAAGAA ACAGTTGCCA
AGCTGTCAAA AGAAATCACA GAACAATTTA GAGTTACAAC
15 AACCATTTCGT GCCTGGAAAA TTAGTATCAC AAGATAATGG
AAAACAAGTT TTACAGACAA GAAAACAAAA GGGTAGCACT
GGTAGTAGTG AAGTTATGGC AAAGAGTGTA TCGAAACCTG
TCCGTGATGG AACAAATTTT CAACAGAAGC AGCAGCAGCA
GCAGCAGCAG CAGCAGCCAC AGTCTAACCA AGAAAAGTTG
20 AATAAGAAAG GTTTGAAAAA AGGTACTAAT ACAGACGATG
TGGTGGGGGT AGAAAGAAAT TTGGCTGAAT CCAATTTTCGT
TAAGGAATAC AACAATCGAA GCCCGGATCC CAAGCTTCCC GGGTACCGC

AGC1 reverse compliment: [529bp]

25 GCGGTACCCG GGAAGCTTGG GATCCGGGCT TCGATTGTTG TATTCCTTAA
CGAAATTGGA TTCAGCCAAA TTTCTTTCTA CCCCACCAC ATCGTCTGTA
TTAGTACCTT TTTCAAACC TTTCTTATTC AACTTTTCTT GGTTAGACTG
TGGCTGCTGC TGCTGCTGCT GCTGCTGCTG CTTCTGTTGA AAATTTGTTC
CATCACGGAC AGGTTTCGAT AACTCTTTG CCATAACTTC ACTACTACCA
30 GTGCTACCCT TTTGTTTTCT TGTCTGTAAG ACTTGTTTTT CATTATCTTG
TGATACTAAT TTTCCAGGCA CGAATGGTTG TTGTAAGTCT AAATTGTTCT
GTGATTTCTT TTGACAGCTT GGCAACTGTT TCTTCTTATT ATTATTATCC

ATTCGACTTG TGTTCTGTTT GTATTGAATT ACCACATTCT GTTGCTCATC
TTTTCTTGCC ATGGTCTTAT CCAAGCTTCC CGGGTAAATC CCGCGGCCAT
GGCGGCCGGG AGCATGCGAC GTCGGGCCC

5 **EXAMPLE 8**

This example illustrates the amplicons produced during the amplification of STR locus **AGC 3** with multiplex cocktails comprising primer pairs SEQ ID NO: 15 and SEQ ID NO: 16.

Sequence for **AGC 3** locus:

10 GCGGTACCCGGGAAGCTTGGATCCTGGTAAAATAAAATTCCAACAGTTCACA
AGTACCAAACACAACCTCCCCCTGGAAAAGGGTCAAGATTTTGTCCAAACAAA
CAGTTAAAAATCAAAATATTACTCCCCCTTTTGTATTATCTAAGGGCCAAAGA
TAACAAACATGAAAATATAGTAATATGTCCAACAAAAGCAAAGAAAGAAA
AAAAAACTTAGTCTCTGTAAAGCTTGACCAAGGTGGACAACCTGCTTTGACAT
15 CTTTGTCTGAACTTCCTCCATGGCAGCAAGACGATTGTTCAACCAGCTGAACCT
CATTCTTGACGTCATGGATTTCTGCGGAAGCAGAATTCGAGCTTGCAAC**agcag**
cagcagcaccagcTTTAGGCCATTTTGAACACACCATCAAAGTATTTGAGGGTT
GGAATGTAGGTCCAATGATAGGGGGCTCAAGTGTTTCATGTGATTGGGCCA
CATTCTTTTGGGAAGATAAAACCTTATAGATTAGATTTGGAAATACAAGTTTA
20 AAGGTTGGCTTTTTATCTCTTCGGAAAGAAACAATCTGGTTCAGAATGTGTGA
GGCCAAATCAATTGAAGCTCCAGAGGTGATGCGGTATAAGAATGATGCCACA
TCTTGAGACACTACGGTCTTGTTGGAGT

AGC3F: ATAGTAATAT GTCCAACAAA AGCAAAGAAA GAAAAA [36bp]

25 AGC3R: CAAGTGTTTC ATGTGATTGG GCCAC [25bp]

AGC3F (rev. comp.): TTTTCTTTC TTGCTTTTG TTGGACATAT TACTAT
[36bp]

AGC3R (rev. comp.): GTGGCCCAAT CACATGAAAC ACTTG [25bp]

30

AGC3 array: AGCAGCAGCA GCACCAGC [18bp]

AGC3 motif: (AGC)₆

AGC3 locus: [660bp]

GCGGTACCCG GGAAGCTTGG ATCCTGGTAA AATAAAATTC
5 CAACAGTTCA CAAGTACCAA ACACAACTCC CCCTGGAAAA
GGGTCAAGAT TTTGTCCAAA CAAACAGTTA AAAATCAAAA
TATTACTCCC CCTTTTTGTT TATCTAAGGG CCAAAGATAA CAAACATGAA
AATATAGTAA TATGTCCAAC AAAAGCAAAG AAAGAAAAAA
AAACTTAGTC TCTGTAAAGC TTGACCAAGG TGGACAACTG CTTTGACATC
10 TTTTGCTGAA CTTCTCCAT GGCAGCAAGA CGATTGTTCA CCAGCTGAAC
CTCATTCTTG ACGTCATGGA TTTCTGCGGA AGCAGAATTC GAGCTTGCAA
CAGCAGCAGC AGCACCAGCT TTAGGCCATT TTTGAAACAC
ACCATCAAAG TATTTGAGG GTTGAATGT AGGTCCAATG
ATAGGGGGCT CAAGTGTTTC ATGTGATTGG GCCACATTCT TTTGGGAAGA
15 TAAACCTTA TAGATTAGAT TTGGAAATAC AAGTTTAAAG GTTGGCTTTT
TATCTCTTCG GAAAGAAACA ATCTGGTTCA GAATGTGTGA
GGCAAATCA ATTGAAGCTC CAGAGGTGAT GCGGTATAAG
AATGATGCCA CATCTTGAGA CACTACGGTC TTGTTGGAGT

20 AGC3 reverse compliment: [660bp]

ACTCCAACAA GACCGTAGTG TCTCAAGATG TGGCATCATT CTTATACCGC
ATCACCTCTG GAGCTTCAAT TGATTTGGCC TCACACATTC TGAACCAGAT
TGTTTCTTTC CGAAGAGATA AAAAGCCAAC CTTTAAACTT GTATTTCCAA
ATCTAATCTA TAAGGTTTTA TCTTCCCAA AGAATGTGGC CCAATCACAT
25 GAAACACTTG AGCCCCCTAT CATTGGACCT ACATTCCAAC CCTCGAAATA
CTTTGATGGT GTGTTTCAAA AATGGCCTAA AGCTGGTGCT GCTGCTGCTG
TTGCAAGCTC GAATTCTGCT TCCGCAGAAA TCCATGACGT
CAAGAATGAG GTTCAGCTGG TGAACAATCG TCTTGCTGCC
ATGGAGGAAG TTCAGCAAAA GATGTCAAAG CAGTTGTCCA
30 CCTTGGTCAA GCTTTACAGA GACTAAGTTT TTTTTCTTT CTTTGCTTTT
GTTGGACATA TTAATATATT TTCATGTTTG TTATCTTTGG CCCTTAGATA
AACAAAAAGG GGGAGTAATA TTTGATTTT TAACTGTTTG TTTGGACAAA

ATCTTGACCC TTTTCCAGGG GGAGTTGTGT TTGGTACTTG TGAAGTGTG
GAATTTTATT TTACCAGGAT CCAAGCTTCC CGGGTACCGC

EXAMPLE 9

- 5 This example illustrates the amplicons produced during the amplification of STR locus **AGC 6** with multiplex cocktails comprising primer pairs SEQ ID NO: 17 and SEQ ID NO: 18.

Sequence for **AGC 6** locus:

TACWTGAGCCCGACGTCGCATGCTCCCGGCCGCCATGGCCCGCGGGATTGCG
10 GTACCCGGGAAGCTTGGCAATATACAATCTSAGKTCACCTCTCTGCTTTCCCAA
GCAGCCCTTGTTTGCAAGTATGCTCAAGACCAACGAAGTACCAGCACTGAGG
CTTGAATGCATGAGTAAAATGTAAAGAAGCCTTCTTTCCCTTTCCGCTTCCAC
TTTCCACCACCAAAAACCTGTGCATGGAAGTATGCCTCTATTCCCTGGTTGTCA
GCAGACAAGAACTGAACAGACGTGGCATATGCGCTGTTTCCTTCACCTGC
15 AAGCGCACTGGCAGCAGCAGCAGCCGACATAGCTGAAGATTTTCCTGACTT**ag**
cagcagcagcagcagcTATTGCAGCAGCAGCAGTTGCTGTATTTAACGTATCAGCAA
ATGATTCAATGTAAATCCATGTTGCAAATGCATACCCATTAGTGAACGGCC
ATCGGCTTTCCCCTGGACCAAGCAAACCAGAGCTTTCACCATCAAACCTCAAA
AGTACATGCTGGTCCCTTTGACTCCTTTCCACTAACTGCCTTCTCCAAAGCAA
20 TCATTAAGCGAGCTGACCAAACAGTGCTAAGTGTTCTTGTGATGACTTGAAA
CCATCTATGCAAATCGATGACACTAAGTG

AGC6F: AGACGTGGCA TATGCGCTGT TCCTTCA [27bp]

AGC6R: GCATACCCAT TAGTGAACGG CCATCGGC [28bp]

25

AGC6F (rev. comp.): TGAAGGAACA GCGCATATGC CACGTCT [27bp]

AGC6R (rev. comp.): GCCGATGGCC GTTCACTAAT GGGTATGC [28bp]

AGC6 array: AGCAGCAGCA GCAGCAGC [18bp]

30

AGC6 motif: (AGC)₆

AGC6 locus: [663bp]

TACWTGAGCC CGACGTCGCA TGCTCCCGGC CGCCATGGCC
CGCGGGATTG CGGTACCCGG GAAGCTTGGC AATATACAAT
CTSAGKTCAC TCTCTGCTTT CCCAAGCAGC CCTTGTTTGC AAGTATGCTC
5 AAGACCAACG AAGTACCAGC ACTGAGGCTT GAATGCATGA
GTAAAATGTA AAGAAGCCTT CTTTCCCTTT CCGCTTCCAC TTTCCACCAC
CAAAAACGTG GCATGGAAGT ATGCCTCTAT TCCCTGGTTG TCAGCAGACA
AGAAACTGAA CAGACGTGGC ATATGCGCTG TTCCTTCACC
TGCAAGCGCA CTGGCAGCAG CAGCAGCCGA CATAGCTGAA
10 GATTTTCCTG ACTTAGCAGC AGCAGCAGCA GCTATTGCAG
CAGCAGCAGT TGCTGTATTT AACGTATCAG CAAATGATTG AATGTAAATC
CATGTTGCAA ATGCATACCC ATTAGTGAAC GGCCATCGGC TTTCCCCTGG
ACCAAGCAAA CCAGAGCTTT CACCATCAAA CTCAAAAGTA
CATGCTGGTC CCTTTGACTC CTTTCCACTA ACTGCCTTCT CCAAAGCAAT
15 CATTAAGCGA GCTGACCAAA CAGTGCTAAG TGTTCTTGTG
ATGACTTGAA ACCATCTATG CAAATCGATG AACTAAGTG AGC

AGC6 reverse compliment: [663bp]

GCTCACTTAG TGTCATCGAT TTGCATAGAT GGTTC AAGT
20 CATCACAAGA AACTTAGCA CTGTTTGGTC AGCTCGCTTA ATGATTGCTT
TGGAGAAGGC AGTTAGTGGA AAGGAGTCAA AGGGACCAGC
ATGTACTTTT GAGTTTGATG GTGAAAGCTC TGGTTTGCTT GGTCCAGGGG
AAAGCCGATG GCCGTTCACT AATGGGTATG CATTTGCAAC ATGGATTTAC
ATTGAATCAT TTGCTGATAC GTTAAATACA GCAACTGCTG
25 CTGCTGCAAT AGCTGCTGCT GCTGCTGCTA AGTCAGGAAA ATCTTCAGCT
ATGTCGGCTG CTGCTGCTGC CAGTGCGCTT GCAGGTGAAG
GAACAGCGCA TATGCCACGT CTGTTCAAGT TCTTGTCTGC TGACAACCAG
GGAATAGAGG CATACTTCCA TGCACAGTTT TTGGTGGTGG
AAAGTGGAAG CGGAAAGGGA AAGAAGGCTT CTTTACATTT
30 TACTCATGCA TTCAAGCCTC AGTGCTGGTA CTTCGTTGGT CTTGAGCATA
CTTGCAAACA AGGGCTGCTT GGGAAAGCAG AGAGTGAMCT

SAGATTGTAT ATTGCCAAGC TTCCCGGGTA CCGCAATCCC GCGGGCCATG
GCGGCCGGGA GCATGCGACG TCGGGCTCAW GTA

EXAMPLE 10

- 5 This example illustrates the amplicons produced during the amplification of STR locus
AGC 8 with multiplex cocktails comprising primer pairs SEQ ID NO: 19 and SEQ ID
NO: 20.

Sequence for **AGC 8** locus:

GCGGTACCCGGGAAGCTTGGATCCCAAGATCCCCTACCTCTTTCGTTCTGAGG
10 CACGCCAGAAGATTTAGAAGTATCAATAGCTCCAAATTCAGAAGAGACACCT
CTGTTAACGGCGTGTCTAAGGTTCCCT**TTCCGACACCGGCGACGCACTCGAG**
CTCCATACGAACATATGAAGGTCCTTGTTTCGGCAGACCATTATT**agcagcagcagca**
gcaggaggaggTGCTGTAACAGTTGTTGCGTCTTTCTTCTTAACAGCCGTATTACTT
GTCGACCCGGAAAACATCGGATTAGGAGGAGGGTAAGACGGGGCAAGACCG
15 CCATTGAAGAGCTCTCCACTCATGCTCCTCGCTCCTCTCTGCT**TTCTTTCCCAT**
ATTTTTCATCATCTCTTCGTCGAAATTAGATGTCCTTGGCGTGACGCCTTTC
GATGACTGAAGTGAGTAGACATCAGCGCCGTGAGTTGGTCCACCACCGTAGC
TGTTGGTGTACCCGTGTTTGGGACTAGCGGCCTTACTGGCATTAAACATGGCG
TAAAAATCAGTCTGGTTGAAGCTCGATGCCCTCGGGGTCGGCTCTCGCGAGG
20 ATTGTACAGAGTAGATCCCAAGCTTCCCGGGTACCGC

AGC8F: TTCCGACACC GGCGACGCAC TC [22bp]

AGC8R: TTCTTTCCCA TATTTTTCAT CATCTCTTCG TCGAA [35bp]

- 25 AGC8F (rev. comp.): GAGTGCGTCG CCGGTGTCGG AA [22bp]

AGC8R (rev. comp.): TTCGACGAAG AGATGATGAA AAATATGGGA AAGAA
[35bp]

AGC8 array: AGCAGCAGCA GCAGCAGGAG GAGG [28bp]

- 30 AGC8 motif: (AGC)5 + (AGG)3

AGC8 locus: [620bp]

GCGGTACCCG GGAAGCTTGG ATCCCAAGAT CCCCTACCTC TTTCGTTCTG
AGGCACGCCA GAAGATTTAG AAGTATCAAT AGCTCCAAAT
TCAGAAGAGA CACCTCTGTT AACGGCGTGT CTAAGGTTC CTTCCGACAC
5 CGGCGACGCA CTCGAGCTCC ATACGAACAT ATGAAGGTCC
TTGTTCGGCA GACCATTATT AGCAGCAGCA GCAGCAGGAG
GAGGTGCTGT AACAGTTGTT GCGTCTTTCT TCTTAACAGC CGTATTACTT
GTCGACCCGG AAAACATCGG ATTAGGAGGA GGGTAAGACG
GGGCAAGACC GCCATTGAAG AGCTCTCCAC TCATGCTCCT CGCTCCTCTC
10 TGCTTCTTTC CCATATTTTT CATCATCTCT TCGTCGAAAT TAGATGTCCT
TGGCGTGACG CCTTTCGATG ACTGAAGTGA GTAGACATCA
GCGCCGTGAG TTGGTCCACC ACCGTAGCTG TTGGTGTACC CGTGTTTGGG
ACTAGCGGCC TTA CTGGCAT TAAACATGGC GTAAAAATCA
GTCTGGTTGA AGCTCGATGC CCTCGGGGTC GGCTCTCGCG AGGATTGTAC
15 AGAGTAGATC CCAAGCTTCC CGGGTACCGC

AGC8 reverse compliment: [620bp]

GCGGTACCCG GGAAGCTTGG GATCTACTCT GTACAATCCT
CGCGAGAGCC GACCCCGAGG GCATCGAGCT TCAACCAGAC
20 TGATTTTAC GCCATGTTA ATGCCAGTAA GGCCGCTAGT CCCAAACACG
GGTACACCAA CAGCTACGGT GGTGGACCAA CTCACGGCGC
TGATGTCTAC TCACTTCAGT CATCGAAAGG CGTCACGCCA AGGACATCTA
ATTTGACGA AGAGATGATG AAAAATATGG GAAAGAAGCA
GAGAGGAGCG AGGAGCATGA GTGGAGAGCT CTTCAATGGC
25 GGTCTTGCCC CGTCTTACCC TCCTCCTAAT CCGATGTTTT CCGGGTCGAC
AAGTAATACG GCTGTTAAGA AGAAAGACGC AACAACTGTT
ACAGCACCTC CTCCTGCTGC TGCTGCTGCT AATAATGGTC TGCCGAACAA
GGACCTTCAT ATGTTTCGTAT GGAGCTCGAG TCGTCGCCG
GTGTCGGAAG GGAACCTTAG ACACGCCGTT AACAGAGGTG
30 TCTCTTCTGA ATTTGGAGCT ATTGATACTT CTAAATCTTC TGGCGTGCCT
CAGAACGAAA GAGGTAGGGG ATCTTGGGAT CCAAGCTTCC
CGGGTACCGC

EXAMPLE 11

This example illustrates the amplicons produced during the amplification of STR locus
AGC 9 with multiplex cocktails comprising primer pairs SEQ ID NO: 21 and SEQ ID
5 NO: 22.

Sequence for AGC 9 locus:

GCGGTACCCGGGAAGCTTGGTACACTCTACATGGCTCAAATTCTCCCGGTAA
GTTGATACATTCCTTCCCAGCATGGAAAACAGAGTAGCCcagcagcagcagcagcag
cagcagcAGTCATATCAATCCAATTGCATTGTATTCTCCTTTAACTCATACAGCT
10 ATAGTTATGGCTGCCAACATATCTTCTCATCTCTTCCACTTAGCTTAATCAACT
CTCTTGGATACTAGGCAATTTCGGTAACAGTTTACAAGTGTTAACCAGACGAC
AAAAAAAGAATTGTACACGTCCAGAATGGTGTCAGGGCCTACTAAAGGTTGA
ACCCAATTATTTTCTCAGGAATGGCTTTTGGCAAACAAGTAGCCTTTGGTCA
CTGCCATTCTGAAGATCCCAAGCTTCCCGGGTACCGC

15

AGC9F: GGTAAGTTGA TACATTCCTT CCC [23bp]

AGC9R: CAAGTAGCCT TTGGTCACTG C [21bp]

AGC9F (rev. comp.): GGGAAGGAAT GTATCAACTT ACC [23bp]

20 AGC9R (rev. comp.): GCAGTGACCA AAGGCTACTT G [21bp]

AGC9 array: AGCAGCAGCA GCAGCAGCAG CAGC [24bp]

AGC9 motif: (AGCC)8

25 AGC9 locus: [411bp]

GCGGTACCCG GGAAGCTTGG TACACTCTAC ATGGCTCAAA
TTCTCCCGGT AAGTTGATAC ATTCCTTCCC AGCATGGAAA ACAGAGTAGC
CAGCAGCAGC AGCAGCAGCA GCAGCACGTC ATATCAATCC
AATTGCATTG TATTCTCCTT TAACTCATAC AGCTATAGTT ATGGCTGCCA
30 ACATATCTTC TCATCTCTTC CACTTAGCTT AATCAACTCT CTTGGATACT
AGGCAATTCTG GTAACAGTTT ACAAGTGTTA ACCAGACGAC

AAAAAAAGAA TTGTACACGT CCAGAATGGT GTCAGGGCCT
 ACTAAAGGTT GAACCCAATT ATTTTCTCAG GAATGGCTTT TGGCAAACAA
 GTAGCCTTTG GTCACTGCCA TTCTGAAGAT CCAAGCTTC CCGGGTACCG
 C

5 AGC9 reverse compliment: [411bp]

GCGGTACCCG GGAAGCTTGG GATCTTCAGA ATGGCAGTGA
 CCAAAGGCTA CTTGTTTGCC AAAAGCCATT CCTGAGAAAA TAATTGGGTT
 CAACCTTTAG TAGGCCCTGA CACCATTCTG GACGTGTACA ATTCTTTTTT
 TGTCGTCTGG TTAACACTTG TAAACTGTTA CCGAATTGCC TAGTATCCAA
 10 GAGAGTTGAT TAAGCTAAGT GGAAGAGATG AGAAGATATG
 TTGGCAGCCA TAACTATAGC TGTATGAGTT AAAGGAGAAT
 ACAATGCAAT TGGATTGATA TGACGTGCTG CTGCTGCTGC TGCTGCTGCT
 GGCTACTCTG TTTTCCATGC TGGGAAGGAA TGTATCAACT TACCGGGAGA
 ATTTGAGCCA TGTAGAGTGT ACCAAGCTTC CCGGGTACCG C

EXAMPLE 12

This example illustrates the amplicons produced during the amplification of STR locus **AGC 10** with multiplex cocktails comprising primer pairs SEQ ID NO: 23 and SEQ ID NO: 24.

20. Sequence for AGC 10 locus:

GCGGTACCCGGGAAGCTTTGGATCAGCGGCAACAACAcagcaacaacatcagca
gcagcagcaacaacaacatcagcagcagcagcagcagcagcagcagcatcaacatcagcaacagcagca
acagcagcagcagcagcagcagcagcagcaacagcagcagcaacagcagcagcaacaacaccagcatcagcaacacca
gcagcagcaaacaccagcatcagcagcaacatcagcagcagcagcTTCAACCGTCACAACAATTGCA
25 TCAGTTGTCTGTT CAGCAGCAGATT CCTAATTGTTATGTCTGCTCTACCCAGT
TTTTCCTCTGGTACTCAGTCTCAGTCTCCATCGCTGCAGGCCATCCCTTCACA
GTGCCAGCAGCCAAGCTTCCCGGGTACCGC

AGC10F: GGATCAGCGG CAACAACAA [19bp]

30 AGC10R: TGTTATGTCT GCTCTACCCA GTTTT [25bp]

AGC10F (rev. comp.): TTGTTGTTGC CGCTGATCC [19bp]

AGC10R (rev. comp.): AAAACTGGGT AGAGCAGACA TAACA [25bp]

AGC10 array: AGCAACAACA ACATCAGCAG CAGCAGCAAC AACAACAACA

5 TCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GCATCAACAT
CAGCAACAGC AGCAACAGCA GCAGCAGCAG CAGCAGCAGC
AACAGCAGCA GCAACAGCAG CAGCAACAAC ACCAGCATCA
GCAACACCAG CAGCAGCAAC ACCAGCATCA GCAGCAACAT
CAGCAGCAGC AGC [213bp]

10

AGC10 motif: (AGC)1 + (AAC)3 + (ATC)1 + (AGC)4 + (AAC)4 + (ATC)1 + (AGC)10
+ (ATC)1 + (AACATC)1 + (AGCAAC)1 + (AGC)2 + (AAC)1 + (AGC)8 + (AAC)1 +
(AGC)3 + (AAC)1 + (AGC)3 + (AAC)2 + (ACC)1 + (AGC)1 + (ATC)1 + (AGC)1 +
(AACACC)1 + (AGC)3 + (AACACC)1 + (AGC)3 + (AACACC)1 + (AGC)3 +
15 (AACACC)1 + (AGCATC)1 + (AGC)2 + (AACATC)1 + (AGC)4

AGC10 locus: [408bp]

GCGGTACCCG GGAAGCTTGG ATCAGCGGCA ACAACAACAG
CAACAACAAC ATCAGCAGCA GCAGCAACAA CAACAACATC
20 AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC ATCAACATCA
GCAACAGCAG CAACAGCAGC AGCAGCAGCA GCAGCAGCAA
CAGCAGCAGC AACAGCAGCA GCAACAACAC CAGCATCAGC
AACACCAGCA GCAGCAACAC CAGCATCAGC AGCAACATCA
GCAGCAGCAG CTTCAACCGT CACAACAATT GCATCAGTTG TCTGTTTCAGC
25 AGCAGATTCC TAATGTTATG TCTGCTCTAC CCAGTTTTTC CTCTGGTACT
CAGTCTCAGT CTCCATCGCT GCAGGCCATC CCTTCACAGT GCCAGCAGCC
AAGCTTCCCG GGTACCGC

AGC10 reverse compliment: [408bp]

30 GCGGTACCCG GGAAGCTTGG CTGCTGGCAC TGTGAAGGGA
TGGCCTGCAG CGATGGAGAC TGAGACTGAG TACCAGAGGA
AAAACTGGGT AGAGCAGACA TAACATTAGG AATCTGCTGC

TGAACAGACA ACTGATGCAA TTGTTGTGAC GGTGAAGCT
GCTGCTGCTG ATGTTGCTGC TGATGCTGGT GTTGCTGCTG CTGGTGTGCTG
TGATGCTGGT GTTGTGCTG CTGCTGTTGC TGCTGCTGTT GCTGCTGCTG
CTGCTGCTGC TGCTGTTGCT GCTGTTGCTG ATGTTGATGC TGCTGCTGCT
5 GCTGCTGCTG CTGCTGCTGA TGTGTTGTT GTTGCTGCTG CTGCTGATGT
TGTGTTGCT GTTGTGTTG CCGCTGATCC AAGCTTCCCG GTTACCGC

EXAMPLE 13

This example illustrates the amplicons produced during the amplification of STR
10 locus **ACT 1** with multiplex cocktails comprising primer pairs SEQ ID NO: 25 and SEQ
ID NO: 26.

Sequence for **ACT 1** locus:

GCGGTACCCGGAAGCTTGGGATCAAAAAACGAGAAGAATATTCATCATGA
AAACTCTATAGAACTTTTATTATTCAAAGTAGGAAGGAACAAGGAAGAGGG
15 AAGAAAAAAAAAAGAAGGGGGCAGAGGGGGGCAATTTATGTTTGCCTTTTATG
CTATATATTTTAGTATCTAGAAGAACAAGAAAAAAGACTATACTCCTAATA
TGAATATGGAATAAAAAATTGACTCAGCATATTAAAGCAGAAACTTTGAA
ATAGACGAACCATGTTTTGGTTTACAACGTGTTGTTTGTATTGACATCTAGT
TGTAAGGAaactactactactACCTGTGCAAAGGTGAACTCTCTACCATGAAAGT
20 AGTAATGGTTTTCAAGGGCCATTAACTTGAACCACCATAGCTAGCAAAGGT
GTTTACATATTCCACTTGTTTGTGAGCCACGCAAAGTGAGTTCCTATTAA
CCAGTTTTTAAACATATGTCATTTCCAAGATAGTTGAAAACCTCGGAAGCAG
CAGCATTACTGTTTTTCATAGCATTTCAGGATTGTTGAAAACCTCAGCAGCA
GCAGCAGCAGCAACAGTATTACTGTTTTTTATAGCATCTCCATTTTGGTTCAC
25 AGTGAAATCCACAGTAAAGGAATTTAGACT

ACT1F: GACTCAGCAT ATTAAAGCAG AACT [25bp]

ACT1R: GTTTACATAT TCCACTTGTT TGTGA [25bp]

30 ACT1F (rev. comp.): AGTTTCTGCT TTAATATGCT GAGTC [25bp]

ACT1R (rev. comp.): TCACAAACAA GTGGAATATG TAAAC [25bp]

ACT1 array: ACTACTACTA CTACT [15bp]

ACT1 motif: (ACT)5

5 ACT1 locus: [660bp]

GCGGTACCCG GGAAGCTTGG GATCAAAAAA CGAGAAGAAT
ATTCATCATG AAAA ACTCTA TAGAACTTTT ATTATTCAAA GTAGGAAGGA
ACAAGGAAGA GGGAAGAAAA AAAAAGAAGG GGGCAGAGGG
GGGCAATTTA TGTTTGCCTT TTATGCTATA TATTTTAGTA TCTAGAAGAA
10 CAAGAAAAAA AGACTATACT CCTAATATGA ATATGGAACT
AAAAAATTGA CTCAGCATAT TAAAGCAGAA ACTTTGAAAT
AGACGAACCA TGTTTTGGTT TACAACGTGT GTTTTTGTAT TGACATCTAG
TTGTAAGGAA CTACTACTAC TACTACCTGT GCAAAGGTG AACTCTCTAC
CATGAAAGTA GTAATGGTTT TCAAGGGCCA TTAACTTGA
15 ACCACCATAG CTAGCAAAGG TGGTTTACAT ATTCCAATTG TTTGTGAGCC
ACGCAAAGTG AGTTCCTATT AACCAGTTTT AAAACATATG TCATTTCCAA
GATAGTTGAA AACCTCGGAA GCAGCAGCAT TACTGTTTTT CATAGCATTT
CCAGGATTGT TGAAAACCTC AGCAGCAGCA GCAGCAGCAA
CAGTATTACT GTTTTTTATA GCATCTCCAT TTTGGTTCAC AGTGAAATCC
20 ACAGTAAAGG AATTTAGACT

ACT1 reverse compliment: [660bp]

AGTCTAAATT CCTTTACTGT GGATTTCACT GTGAACCAAA ATGGAGATGC
TATAAAAAAC AGTAATACTG TTGCTGCTGC TGCTGCTGCT GAAGTTTTCA
25 ACAATCCTGG AAATGCTATG AAAAACAGTA ATGCTGCTGC
TTCCGAGGTT TTCAACTATC TTGGAAATGA CATATGTTTT AAAACTGGTT
AATAGGAACT CACTTTGCGT GGCTCACAAA CAAGTGGAAT
ATGTAAACCA CCTTTGCTAG CTATGGTGGT TCAAGTTAAA TGGCCCTTGA
AAACCATTAC TACTTTCATG GTAGAGAGTT CACCTTTTGC ACAGGTAGTA
30 GTAGTAGTAG TTCCTTACAA CTAGATGTCA ATACAAAAAC
CACAGTTGTA AACCAAAACA TGGTTCGTCT ATTTCAAAGT TTCTGCTTTA
ATATGCTGAG TCAATTTTTT AGTTCCATAT TCATATTAGG AGTATAGTCT

TTTTTTCTTG TTCTTCTAGA TACTAAAATA TATAGCATAA AAGGCAAACA
TAAATTGCCC CCCTCTGCCC CCTTCTTTTT TTTTCTTCCC TCTTCCTTGT
TCCTTCCTAC TTTGAATAAT AAAAGTTCTA TAGAGTTTTT CATGATGAAT
ATTCTTCTCG TTTTTTGATC CCAAGCTTCC CGGGTACCGC

5

EXAMPLE 14

This example illustrates the amplicons produced during the amplification of STR locus **CCT 2** with multiplex cocktails comprising primer pairs SEQ ID NO: 27 and SEQ ID NO: 28.

10 Sequence for **CCT 2** locus:

GCGGTACCCGGAAGCTTGGGATCGT**GCAGTGGATGTGTCTCGGGTTCGAAA**
GTCTAT**cctcctcctcctcct**GCCGTTGGA
ATGGTGTGTTTCGTCTCTGCCTGTTCAAAGAGCGACAATCAATGGTCTTAAAGG
AGCACCTATCTGCCTGACTGGAAATCCAAGCTCCCTCCGATGAATGATTGTTT
15 GTTCTTGCTTGATTACCGGAGGACCGACGCAGGAAGGCGTTGTCACTGCGAC
TTGGTGCCTACTATGCTCTTCACGGAAAGGAGTGAAACGAGCAAGGAGAGAG
TCAACCTTAATGTCAGTGATAATAGTAAAGGAAGAGACAGAATCTCATCTGC
TTGGCTGGTCGACACAAGCAATGCCCAAAGAGCATTCTTTTCTATTTTCATGC
TTCATAATGTATCCGCCGGATTGAAACAGTCTCT**TTTGTGCCTGACCTAATC**
20 **CTCTAG**CTCTTTACTTGCCAGGAGAAGGCTCGCCAAGCTTCCCGGGTACCGC

CCT2F: GCAGTGGATG TGTCGGGT [18bp]

CCT2R: TTTGTGCCTG ACCTAATCCT CTA [23bp]

25 CCT2F (rev. comp.): ACCCGACACA TCCACTGC [18bp]

CCT2R (rev. comp.): TAGAGGATTA GGTGAGGCAC AAA [23bp]

CCT2 array: CCTCCTCCTC CTCCT [15bp]

CCT2 motif: (CCT)5

30

CCT 2 locus: [499bp]

GCGGTACCCG GGAAGCTTGG GATCGTGCAG TGGATGTGTC
GGGTTCGAAA GTCTATCCTC CTCCTCCTCC TGCCGTTGGA ATGGTGTGTT
CGTCTCTGCC TGTTC AAAGA GCGACAATCA ATGGTCTTAA
AGGAGCACCT ATCTGCCTGA CTGGAAATCC AAGCTCCCTC
5 CGATGAATGA TTGTTTGTTT TTGCTTGATT ACCGGAGGAC CGACGCAGGA
AGGCGTTGTC ACTGCGACTT GGTGCCTACT ATGCTCTTCA CGGAAAGGAG
TGAAACGAGC AAGGAGAGAG TCAACCTTAA TGTCAGTGAT
AATAGTAAAG GAAGAGACAG AATCTCATCT GCTTGGCTGG
TCGACACAAG CAATGCCCAA AGAGCATTCT TTTCTATTTT CATGCTTCAT
10 AATGTATCCG CCGGATTGAA ACAGTCTCTT TTGTGCCTGA CCTAATCCTC
TAGCTCTTTA CTTGCCAGGA GAAGGCTCGC CAAGCTTCCC GGGTACCGC

CCT 2 locus reverse compliment: [499bp]

GCGGTACCCG GGAAGCTTGG CGAGCCTTCT CCTGGCAAGT
15 AAAGAGCTAG AGGATTAGGT CAGGCACAAA AGAGACTGTT
TCAATCCGGC GGATACATTA TGAAGCATGA AAATAGAAAA
GAATGCTCTT TGGGCATTGC TTGTGTCGAC CAGCCAAGCA GATGAGATTC
TGTCTCTTCC TTTACTATTA TCACTGACAT TAAGGTTGAC TCTCTCCTTG
CTCGTTTCAC TCCTTTCCGT GAAGAGCATA GTAGGCACCA AGTCGCAGTG
20 ACAACGCCTT CCTGCGTCGG TCCTCCGTA ATCAAGCAAG
AACAAACAAT CATTATCGG AGGGAGCTTG GATTTCAGT
CAGGCAGATA GGTGCTCCTT TAAGACCATT GATTGTCGCT CTTTGAACAG
GCAGAGACGA ACACACCATT CCAACGGCAG GAGGAGGAGG
AGGATAGACT TTCGAACCCG ACACATCCAC TGCACGATCC
25 CAAGCTTCCC GGGTACCGC

[0052] While certain of the preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following
30 claims.

Table 1.

Collection of worldwide samples with representatives from all continents except Australia.

Continent

<u>North America</u>	<u># of Samples</u>
U.S.A.	188
Canada	1
Mexico	7

Total North America 196

<u>Central & South America</u>	<u># of Samples</u>
Colombia	3
Costa Rica	6
Jamaica	4

Total C & S America 13

<u>Africa</u>	<u># of Samples</u>
Nigeria	1
South Africa	6
Sierra Leone	2
Uganda	2
Zimbabwe	2

Total Africa 13

<u>Asia</u>	<u># of Samples</u>
Afghanistan	14
Cambodia	1
China	4
India	5
Japan	3
Korea	4
Kurdistan	2
Nepal	1
Pakistan	2
Russia	4
Thailand	1
Turkey	3
Uzbekistan	2

Total Asia 46

Table 1. Continued...

Europe	# of Samples
Czechoslovakia	1
France	3
Germany	4
Holland	2
Hungary	8
Italy	3
Poland	3
Romania	1
Spain	2
Total Europe	27

Total # Samples = 295

Table 2. Attributes of eight microsatellite loci developed for *Cannabis sativa*. Values in the 'Amplicon Size Range (bp)' refer to results from fragment analyses of 295 *C. sativa* samples. 'Number of Alleles' reflects the number of alleles observed in this data set.

Locus Name Dye Label ^a	Primer Sequences	Repeat Motifs ^b	Amplicon Size Range (bp)	T _m (°C)	Number of Alleles	H _E
AAAG1 HEX	F: 5' GTCAGAAAGCGAAGACCTTTAGA 3' R: 5' GATGATGGCTGCCTGCTTAC 3'	(AAAG) ₆	103-135	59	16	0.684
AAAG5 NED	F: 5' GTCAATTAAATGCTTATAGCCATATGTTTCTACTAC 3' R: 5' GCAACTTCAGGAATACCTTGTCTCTCTGTTCT 3'	(AAAG) ₅	188-200	59	4	0.625
AGC1 FAM	F: 5' GCAAAAGAGTGATGCAAAACCTGTC 3' R: 5' GCCCAACCACTGCTGTGATAGTAC 3'	(AGC) ₁₀	128-164	59	10	0.656
AGC6 HEX	F: 5' GAGACGTCCCATATGGCTGTTCCTTCA 3' R: 5' GCCGATGCCGTTCACTAATGGGTATGCC 3'	(AGC) ₆	200 & 221	62	2	0.132
AGC8 NED & FAM	F: 5' GTTCCGACACCCGGCGACCCACATC 3' R: 5' GTTCGACGACAGATGATGAATAATGCGAAAGAA 3'	(AGC) ₅	264-279	59	6	0.591
AGC9 HEX	F: 5' GGTAAAGTTGATACATTCCTTCCC 3' R: 5' GCAGTGACCAACCCCTACTTG 3'	(AGC) ₉	317-335	62	7	0.698
AGC10 NED	F: 5' GGATCAGCCGCAACAACAA 3' R: 5' GAATACTGGGTAGACGACACATAACA 3'	(AGC) ₄₃	273-327	62	15	0.776
ACT1 FAM	F: 5' GACTCAGCATATTAAAGCAGAAACT 3' R: 5' GTCACAAACAAGTGGAAATATGTAAAC 3'	(ACT) ₆	218-224	59	3	0.440

^aHEX & FAM labeled primers were ordered from Integrative DNA Technologies; NED labeled primers were ordered from Perkin Elmer

^bMost repeat motifs are not perfect and appear to be complete

APPENDIX 1: Raw STR Data

Allelic scores, in base pairs, for all 295 samples genotyped across eight polymorphic loci. Samples where the same allelic size is listed twice are homozygous, whereas two different allelic sizes indicate a heterozygous state. Marker names are displayed across the top row of each page.

LEGEND

AFG	Afghanistan	UGA	Uganda
AK	Alaska, USA	UZB	Uzbekistan
AZ	Arizona, USA	WV	West Virginia, USA
CA	California, USA	ZIM	Zimbabwe
CAM	Cambodia			
CAN	Canada			
CHI	China			
COL	Colombia			
CoR	Costa Rica			
CT	Connecticut, USA			
CZE	Czechoslovakia			
FRA	France			
GER	Germany			
HA	Hawaii, USA			
HOL	Holland			
HUN	Hungary			
IND	India			
ITA	Italy			
JAM	Jamaica			
JAP	Japan			
KOR	Korea			
KURD	Kurdistan			
KY	Kentucky, USA			
MEX	Mexico			
NEP	Nepal			
NIG	Nigeria			
OR	Oregon, USA			
PAK	Pakistan			
POL	Poland			
ROM	Romania			
RUS	Russia			
SAF	South Africa			
SLe	Sierra Leone			
SPA	Spain			
THI	Thailand			
TN	Tennessee, USA			
TUR	Turkey			

Sample	A A G 1		A C T 1		A G C 8		A G C 9		A G C 1		A A A G 5		A G C 6		A G C 10	
AFG177	127	127	221	221	264	270	326	326	152	152	192	192	200	200	309	321
AFG178	127	127	221	221	264	270	326	326	140	152	192	192	200	200	321	321
AFG181	103	117	221	221	270	270	326	326	140	152	196	196	200	200	294	303
AFG182	103	117	221	221	270	270	326	326	140	152	196	196	200	200	306	306
AFG217	117	117	218	221	264	264	326	326	152	164	192	192	200	200	303	309
AFG218	117	123	218	218	264	270	326	326	152	152	192	192	200	200	309	309
AFG223	117	127	218	221	270	270	332	332	131	131	188	196	200	200	309	309
AFG224	117	127	218	221	270	270	320	326	137	152	196	196	200	200	300	300
AFG225	127	127	221	221	267	267	320	320	140	152	188	192	200	200	300	309
AFG61	123	127	221	221	270	270	326	326	164	164	188	192	200	200	300	309
AFG62	117	127	218	221	270	270	326	326	131	131	192	192	200	221	300	300
AFG63	117	117	218	218	270	270	326	326	152	164	192	192	200	200	300	309
AFG64	127	127	218	218	270	270	326	326	152	164	192	192	200	200	309	324
AFG83	127	127	221	221	270	276	326	326	152	152	196	196	200	200	309	312
AK81	117	127	218	221	270	270	326	326	152	152	188	188	200	200	300	300
AK82	117	127	221	221	270	270	329	332	152	152	188	188	200	200	300	315
AZ100	117	127	218	218	270	276	323	332	131	152	196	200	200	221	309	315
AZ101	117	127	218	221	270	276	326	326	131	131	188	200	200	221	309	315
AZ102	117	123	221	221	264	270	323	332	131	152	188	200	200	200	315	315
AZ103	117	127	218	218	270	270	326	332	140	140	188	192	200	200	315	324
AZ104	117	117	218	221	270	276	323	332	131	152	192	200	200	200	309	315
AZ176	127	127	218	218	264	264	326	332	140	140	192	196	200	200	300	309
AZ97	117	123	221	221	264	270	323	332	131	152	188	200	200	200	315	315
AZ98	117	117	221	221	270	270	323	332	131	152	192	200	200	200	315	315
AZ99	127	127	218	221	264	276	326	326	152	152	192	196	221	221	309	309
CA121	117	127	221	221	264	276	323	329	137	152	188	192	200	200	315	321
CA122	117	117	221	221	264	270	326	326	137	152	192	192	200	200	312	321
CA123	117	117	221	221	264	270	323	323	137	152	192	192	200	200	309	315
CA124	121	123	221	221	264	264	323	326	152	152	192	192	200	200	306	309
CA125	123	127	218	221	264	270	326	326	152	152	192	192	200	200	306	306
CA126	127	127	218	221	264	270	326	332	131	152	192	196	200	200	312	315
CA127	127	127	218	221	264	270	326	332	131	152	192	196	200	200	312	315
CA128	117	117	224	224	270	270	326	326	140	152	188	188	200	200	300	300

Sample	AAAG1	ACT1	AGC8	AGC9	AGC1	AAAG5	AGC6	AGC10
CA129	117	221	270	326	140	188	200	300
CA130	113	221	264	326	152	192	200	309
CA131	113	221	264	326	152	188	200	309
CA132	127	218	270	326	152	188	200	309
CA133	113	221	264	329	152	192	200	315
CA134	123	221	270	326	152	192	200	309
CA135	127	221	264	326	152	192	200	309
CA136	117	221	264	326	152	188	200	309
CA137	117	221	270	323	152	192	200	312
CA138	117	221	270	326	152	188	200	309
CA139	117	221	264	326	152	192	200	309
CA140	117	221	264	323	146	192	200	315
CA141	117	221	264	323	152	192	200	309
CA142	117	221	264	323	152	192	200	309
CA143	117	221	264	326	152	188	200	309
CA144	117	221	264	326	152	188	200	315
CA145	117	221	264	323	137	192	200	309
CA146	117	221	264	323	137	192	200	309
CA147	117	221	264	326	152	188	200	315
CA148	117	221	270	326	140	192	200	309
CA149	127	221	264	326	152	188	200	318
CA150	117	221	264	320	152	192	200	288
CA151	117	221	270	323	152	188	200	300
CA152	117	221	270	326	152	188	200	309
CA153	123	221	264	326	146	192	200	309
CA154	117	221	264	320	137	192	200	297
CA155	107	221	264	326	137	192	200	297
CA156	117	221	270	326	137	192	200	297
CA157	117	221	270	326	137	192	200	303
CA158	111	221	270	320	146	196	200	297
CA159	117	221	264	323	152	188	200	309
CA160	117	221	264	326	131	188	200	303
CA161	117	221	270	323	152	188	200	309

Sample	AAAG1	ACT1	AGC8	AGC9	AGC1	AAAG5	AGC6	AGC10
CoR170	117	218	279	323	164	188	200	309
CoR171	117	221	270	323	164	188	200	309
CoR172	117	218	270	323	146	192	200	309
CoR173	117	218	264	323	146	188	200	309
CoR174	117	218	270	323	152	188	200	309
CoR175	117	218	270	323	152	188	200	309
CT1	117	221	264	323	140	188	200	309
CT2	117	221	264	323	140	188	200	309
CT3	117	221	264	326	140	188	200	300
CT4	117	218	264	326	152	188	200	309
CT5	117	221	264	326	140	188	200	315
CT6	117	221	264	326	140	188	200	315
CT7	117	221	264	326	140	188	200	300
CT8	117	221	264	326	140	188	200	300
CT9	117	218	264	326	140	188	200	309
CT10	117	221	264	326	152	188	221	300
CT11	117	218	270	326	140	192	200	309
CT12	117	221	264	332	140	188	200	309
CT13	117	221	264	332	140	188	200	309
CT14	123	221	270	326	140	188	200	309
CT15	117	221	264	326	152	188	221	300
CT16	117	221	264	326	140	188	200	300
CT17	117	221	264	326	140	188	200	300
CT18	117	221	264	326	152	188	221	300
CT19	123	221	270	326	140	188	200	309
CT20	117	221	264	326	152	188	221	300
CT21	117	218	264	326	140	188	200	309
CT22	117	221	264	326	131	188	200	321
CT23	117	221	264	323	131	192	200	309
CT24	117	221	270	326	152	188	200	309
CT25	117	221	264	326	152	196	200	321
CT26	117	221	264	323	140	188	200	309
CT27	117	221	264	326	140	188	200	300

Sample	AAAG1	ACT1	AGC8	AGC9	AGC1	AAAG5	AGC6	AGC10
CT 28	117	221	264	323	140	188	200	309
CT 29	127	221	264	326	131	188	200	321
CT 30	117	221	264	326	152	188	200	309
CT 31	117	221	270	323	140	188	200	309
CT 32	117	221	264	323	140	188	200	309
CT 33	117	221	270	326	152	188	200	300
CT 34	117	221	270	323	140	188	200	309
CT 35	117	218	264	326	152	188	200	309
CT 36	117	221	264	326	131	188	200	321
CT 37	127	221	264	323	131	192	200	321
CT 38	117	221	276	323	140	188	200	309
CT 39	117	221	276	332	152	188	200	309
CT 40	117	221	264	326	131	188	200	321
CZE187	117	221	270	329	146	192	200	303
FRA189	103	218	264	332	134	192	200	294
FRA190	113	218	264	320	146	192	200	306
FRA193	117	221	264	332	134	192	200	318
GER188	117	218	264	332	146	188	200	312
GER195	117	218	264	329	128	192	200	303
GER240	115	221	264	326	146	192	200	294
GER91	103	221	279	320	146	192	200	321
HA209	117	221	264	326	152	188	200	315
HA210	117	221	264	326	152	188	200	315
HA211	117	221	270	326	131	188	200	324
HA77	117	218	264	326	137	188	200	297
HA78	117	221	264	329	152	188	200	315
HA79	117	221	264	326	152	188	200	300
HA80	117	221	264	326	152	188	200	300
HOL200	123	218	264	326	152	192	200	312
HOL230	117	221	264	329	140	188	200	309
HUN192	117	218	270	329	146	188	200	321
HUN198	117	221	270	326	146	192	200	318
HUN212	105	218	270	329	134	188	200	294

Sample	AAG1		ACT1		AGC8		AGC9		AGC1		AAG5		AGC6		AGC10	
HUN213	115	117	218	221	264	270	329	332	137	146	188	192	200	200	303	303
HUN70	117	117	218	218	270	270	332	332	146	146	188	192	200	200	315	321
HUN84	117	117	218	221	264	264	326	329	146	152	192	192	200	200	303	303
HUN87	117	121	218	221	264	270	326	329	137	152	188	188	200	200	273	303
HUN89	117	117	221	221	264	279	326	329	128	146	192	192	200	200	303	306
IND179	123	123	221	221	270	276	323	329	140	152	192	196	200	200	303	303
IND180	113	127	221	221	270	270	326	332	152	152	192	196	200	200	306	309
IND207	121	123	218	221	270	270	323	323	152	152	192	196	200	200	309	309
IND229	123	123	218	221	276	279	326	326	152	152	188	192	200	200	306	306
IND86	117	117	218	221	279	279	326	326	152	164	192	192	200	200	309	309
ITA191	117	117	218	221	270	270	317	329	146	152	192	192	200	200	297	306
ITA194	121	121	218	218	270	270	332	332	134	143	188	192	200	200	300	318
ITA88	103	117	218	218	264	270	320	329	146	152	192	192	200	200	306	306
JAM236	117	117	221	221	270	279	329	329	164	164	188	188	200	200	300	300
JAM237	117	117	221	221	264	270	329	329	164	164	188	188	200	200	309	309
JAM65	117	123	218	218	270	276	320	326	152	164	196	200	200	221	300	321
JAM66	127	127	218	221	270	270	326	329	152	164	188	200	200	200	309	309
JAP196	113	113	218	218	270	270	326	326	143	146	196	196	200	200	306	306
JAP241	109	123	221	221	270	270	320	320	128	128	192	200	200	200	306	306
JAP242	103	109	221	221	270	270	317	326	128	143	192	200	200	200	300	306
KOR186	109	113	218	221	270	270	320	326	128	146	192	192	200	200	321	321
KOR204	113	123	218	221	270	270	320	326	134	146	192	196	200	200	297	297
KOR248	113	117	221	221	270	270	326	329	128	128	192	196	200	200	297	306
KOR249	113	123	218	221	270	270	326	326	131	137	192	192	200	200	294	294
KURD214	119	119	221	221	264	264	326	326	128	152	192	192	200	200	294	294
KURD215	117	123	221	221	264	264	326	326	152	152	192	192	200	200	306	306
KY1	125	133	221	221	270	270	326	326	152	152	192	196	200	200	309	309
KY165	117	117	221	221	270	270	326	326	152	152	188	196	200	200	309	309
KY166	117	117	221	221	270	270	326	326	152	152	188	188	200	200	309	309
KY167	117	127	221	221	270	270	326	329	152	152	188	188	200	200	309	321
KY168	117	127	224	224	264	270	326	326	152	152	196	196	200	200	309	321
KY169	117	127	218	221	264	270	323	326	152	152	192	192	200	200	309	309
KY2	123	133	218	221	270	270	323	326	140	152	188	192	200	200	315	318

Sample	AAAG1		ACT1		AGC8		AGC9		AGC1		AAAG3		AGC6		AGC10	
KY25	121	133	218	221	270	270	323	326	152	152	196	196	221	221	306	312
KY26	123	123	221	221	270	270	323	329	152	152	188	188	200	221	303	309
KY27	123	123	218	221	270	270	323	326	152	152	192	192	200	221	303	309
KY28	123	123	221	221	270	270	323	326	137	152	196	196	200	200	303	303
KY29	113	127	218	221	270	270	323	326	137	152	192	192	221	221	306	312
KY3	123	127	218	221	270	270	323	332	137	164	192	192	200	221	312	327
KY30	117	123	218	221	270	270	326	329	137	137	188	196	200	221	309	309
KY31	113	117	218	218	264	270	323	326	140	140	192	192	200	200	306	312
KY32	119	127	218	221	264	264	326	326	152	152	196	196	200	200	309	309
KY4	117	123	221	221	264	270	323	329	140	152	188	196	200	200	306	309
KY49	123	123	221	221	264	270	323	332	140	152	188	188	200	200	312	312
KY5	117	127	218	221	264	264	329	329	137	152	192	192	200	200	306	327
KY50	117	123	221	221	270	270	329	329	152	164	188	188	200	200	303	318
KY51	117	123	221	221	270	270	320	323	152	152	188	188	200	221	309	321
KY52	117	117	218	218	264	270	323	329	137	152	192	192	200	200	303	315
KY53	123	133	221	221	264	273	323	332	137	152	192	192	200	221	309	309
KY54	117	127	221	221	270	270	326	332	140	152	192	196	200	200	309	318
KY55	133	133	221	221	264	264	335	335	152	152	196	196	200	200	309	309
KY56	135	135	221	221	264	264	326	335	152	152	188	188	200	200	303	312
KY6	123	133	221	221	264	273	323	332	152	152	192	192	200	200	309	309
KY7	117	127	221	221	270	279	323	326	152	152	188	192	200	200	312	312
KY74	123	123	221	221	270	270	323	329	152	152	188	192	200	200	297	309
KY75	117	117	218	221	264	273	323	326	152	152	192	196	200	200	309	309
KY76	117	117	218	221	270	270	329	329	137	137	188	192	200	200	315	315
KY8	123	129	218	218	270	270	326	326	152	152	188	188	200	200	303	312
MEX233	117	121	218	218	264	264	320	329	137	152	192	192	200	200	318	318
MEX246	117	117	221	221	264	264	323	323	152	164	188	188	200	221	309	309
MEX57	117	117	218	221	264	270	323	323	131	152	188	192	200	200	309	309
MEX58	117	117	221	221	264	264	326	329	152	164	188	192	200	200	309	309
MEX59	113	117	218	218	270	270	329	332	152	164	192	192	221	221	288	315
MEX60	117	117	218	221	264	279	323	329	164	164	188	188	200	200	309	315
MEX85	117	117	218	221	264	270	323	329	152	164	188	188	200	221	309	315
NEP221	123	123	218	218	270	270	317	323	152	152	188	192	200	200	288	318

Sample	AAAG1		ACT1		AGC8		AGC9		AGC1		AAAG5		AGC6		AGC10	
NIG222	123	123	218	218	264	270	323	326	134	155	192	192	200	200	309	309
OR93	117	117	221	221	264	270	326	332	140	152	188	188	200	200	300	309
OR94	117	117	221	221	264	270	326	332	140	152	188	188	200	200	300	309
PAK226	115	127	218	221	264	270	326	332	152	152	192	196	200	200	300	300
PAK227	117	127	218	218	264	264	326	326	152	152	188	196	200	200	300	300
POL216	117	117	218	221	270	270	332	332	146	146	188	188	200	200	300	318
POL228	121	123	221	221	264	270	332	332	134	152	192	192	200	200	297	297
POL71	121	121	221	221	270	270	332	332	131	152	192	192	200	200	303	303
ROM203	117	117	218	218	264	270	320	320	137	146	188	192	200	200	321	321
RUS197	117	121	218	218	270	270	329	332	146	146	192	192	200	200	300	303
RUS205	117	117	218	218	270	270	326	332	152	152	192	192	200	200	312	318
RUS206	117	117	218	218	270	270	326	332	137	140	192	192	200	200	300	312
RUS90	121	125	218	218	270	270	326	329	146	152	192	192	200	200	297	315
SAF208	115	115	221	221	276	276	326	335	128	152	192	192	200	200	309	309
SAF220	117	117	218	218	264	264	323	323	152	152	192	192	200	200	309	309
SAF247	123	123	221	221	264	264	323	323	152	155	192	192	200	200	309	309
SAF250	117	117	221	221	264	264	323	323	152	152	192	192	200	200	309	309
SAF251	117	123	218	218	264	264	323	326	152	152	192	192	200	200	309	309
SAF252	117	117	221	221	264	264	323	329	152	155	192	192	200	200	309	309
SLe234	123	123	218	218	270	270	323	326	152	155	192	192	200	200	309	309
SLe235	123	123	218	218	270	270	326	326	152	152	192	192	200	200	309	309
SPA202	119	123	221	221	270	270	326	329	128	143	192	192	200	200	306	306
SPA92	115	115	221	221	264	276	329	332	152	152	192	192	200	200	303	303
THI232	123	123	221	221	270	270	326	326	152	152	192	192	200	200	300	309
TN10	117	123	221	221	270	270	320	323	152	152	192	192	200	221	303	309
TN105	123	123	218	221	270	279	323	323	128	128	188	192	200	200	309	309
TN106	123	123	218	221	270	279	323	323	128	128	188	192	200	200	309	309
TN107	127	127	221	221	264	264	326	326	152	152	188	188	200	200	303	303
TN108	117	117	221	221	264	270	326	329	152	152	188	188	200	200	303	312
TN109	117	117	221	221	264	276	323	329	128	146	188	192	200	200	309	309
TN11	117	127	221	221	264	276	326	332	152	164	192	192	200	221	309	309
TN110	117	127	218	221	264	270	326	332	152	164	192	200	200	200	309	315
TN111	117	127	218	221	264	270	326	332	152	164	192	200	200	200	309	315

Sample	A A G I		A C T I		A G C 8		A G C 9		A G C 1		A A G 5		A G C 6		A G C 10	
TN112	115	123	221	221	264	264	329	329	152	152	188	192	200	200	300	321
TN113	117	127	218	221	270	270	326	329	152	152	192	192	200	200	309	309
TN114	117	117	218	218	270	270	326	329	152	152	188	188	200	200	300	303
TN115	117	117	221	221	264	270	317	323	152	152	192	192	221	221	297	309
TN116	117	123	221	221	270	279	323	326	164	164	188	192	200	200	309	309
TN117	117	117	221	221	270	270	323	326	152	152	188	188	200	200	309	309
TN118	117	117	218	221	270	273	326	326	152	152	188	188	200	200	300	309
TN119	117	117	218	221	273	276	329	329	152	152	188	200	200	200	315	315
TN12	117	127	221	221	264	270	326	326	152	152	192	192	200	200	318	318
TN120	117	117	218	218	273	273	329	329	152	152	188	200	200	200	300	315
TN13	117	127	221	221	264	270	326	326	152	152	188	192	200	221	309	318
TN14	117	125	221	221	264	264	326	332	137	137	200	200	200	200	294	318
TN15	117	123	221	221	270	273	329	329	128	152	188	192	200	200	312	315
TN16	117	117	218	221	270	279	329	332	137	137	188	188	200	200	300	315
TN41	115	127	218	221	270	270	320	326	140	152	196	196	200	200	312	324
TN42	115	127	218	221	270	270	320	326	140	152	196	196	200	200	309	321
TN43	113	117	221	221	264	270	320	326	140	155	188	196	200	200	309	309
TN44	127	127	221	221	264	270	320	326	140	155	196	196	200	200	309	321
TN45	117	117	218	221	264	279	323	329	152	152	188	188	200	221	300	309
TN46	117	127	221	221	264	270	326	329	152	152	196	196	200	200	309	321
TN47	127	127	221	221	264	264	326	329	140	164	192	192	200	200	303	309
TN48	117	127	221	221	270	270	323	326	137	137	188	188	200	200	300	309
TN9	117	127	221	221	264	270	323	329	152	152	192	192	200	221	315	315
TUR199	113	117	218	221	264	264	326	326	134	152	192	192	200	200	309	312
TUR253	103	117	218	221	270	270	326	329	146	152	188	188	200	200	303	312
TUR254	115	117	218	218	270	270	326	329	152	152	188	192	200	200	303	309
UGA 238	115	115	218	218	264	264	326	326	152	152	192	192	200	200	309	309
UGA 239	115	115	218	218	264	264	326	326	152	152	192	192	200	200	309	309
UZB255	115	117	218	218	270	270	332	332	137	137	192	192	200	200	300	300
UZB256	123	123	218	221	264	270	323	326	137	152	192	196	200	200	306	306
WV151	113	117	221	221	270	270	323	326	152	164	192	196	221	221	300	309
WV152	117	117	218	221	270	270	323	323	137	152	192	192	200	200	318	318
WV153	127	127	218	221	264	264	326	329	152	152	188	192	200	221	315	321

Sample	A A A G 1		A C T 1		A G C 8		A G C 9		A G C 1		A A A G 5		A G C 6		A G C 10	
W V 154	123	127	218	221	264	264	323	329	152	152	192	196	200	200	309	309
W V 155	123	127	221	221	264	264	326	326	143	164	188	192	200	221	309	309
W V 156	123	127	221	221	264	264	326	326	143	164	188	192	200	221	309	309
W V 157	117	123	218	221	270	270	329	332	152	152	188	192	200	200	309	315
W V 158	117	127	221	221	270	270	320	326	152	152	192	192	200	200	300	321
W V 159	117	127	218	221	270	270	326	326	131	152	192	192	200	200	309	312
W V 160	123	127	218	218	270	270	323	332	152	152	192	192	200	200	309	309
W V 161	123	127	218	218	270	270	323	326	131	131	192	196	200	221	309	309
W V 162	117	123	218	221	264	270	323	326	137	140	192	192	200	200	309	315
W V 163	123	123	221	221	270	270	323	326	137	140	188	192	200	200	309	315
W V 164	123	127	218	218	270	270	326	332	140	152	192	200	200	200	306	309
W V 17	125	125	221	221	270	270	326	329	146	152	188	196	200	200	309	321
W V 18	117	117	221	221	264	270	326	326	140	152	196	196	200	200	300	303
W V 19	117	117	221	221	264	270	323	329	152	164	192	192	200	200	309	315
W V 20	127	127	218	218	270	270	326	326	131	131	192	192	200	200	309	309
W V 21	117	123	221	221	270	276	326	326	140	140	188	196	200	200	300	315
W V 22	117	123	221	221	270	270	323	326	140	140	188	196	200	200	309	309
W V 23	117	127	221	221	270	270	323	323	152	152	188	192	200	200	309	315
W V 24	127	127	221	221	264	276	326	326	152	152	192	192	200	200	309	312
W V 33	117	117	221	221	270	270	329	329	152	152	192	192	200	200	309	315
W V 34	123	123	221	221	264	270	323	326	155	155	188	188	200	221	309	321
W V 35	117	123	218	218	270	270	326	332	152	152	196	200	200	200	309	309
W V 36	123	123	218	221	264	270	323	326	152	152	192	196	200	221	309	312
W V 37	117	123	218	221	270	279	326	326	137	137	188	192	200	221	315	321
W V 38	127	127	218	218	270	270	326	326	131	131	192	192	200	200	309	309
W V 39	117	123	221	221	264	270	329	329	152	164	188	192	200	221	309	309
W V 40	117	127	218	221	270	270	323	326	152	152	192	196	200	200	309	309
W V 95	117	117	221	221	270	270	323	332	164	164	192	196	200	200	309	315
W V 96	117	123	218	221	264	270	329	332	146	152	192	192	200	200	309	315
ZIM 244	117	123	218	218	264	270	323	326	152	152	192	192	200	200	309	309
ZIM 245	123	123	218	218	264	270	326	326	152	152	192	192	200	200	309	309

REFERENCES

- [1] G.S. Chopra, Man and marijuana, *Int. J. Addiction*, 4 (2) (1969) 215–247.
- [2] R.E. Schultes, Man and marijuana, *Nat. Hist.* 82 (7) (1973) 58–63, 80, 82.
- [3] R. Adams, Marijuana, *Science* 92 (1940) 115–119.
- 5 [4] E. Small, H.D. Beckstead, A. Chan, The evolution of cannabinoid phenotypes in *Cannabis*, *Econ. Bot.* 29 (1975) 219–232.
- [5] W.A. Emboden, Jr., Ritual use of *Cannabis sativa* L.: a historical-ethnographic survey, in: P. T. Furst (Eds.), *Flesh of the gods: the ritual use of hallucinogens*, Praeger Publishers, NY and WA, 1972, pp. 214–235.
- 10 [6] E. Small, A. Cronquist, A practical and natural taxonomy for *Cannabis*, *Taxon* 25 (4) (1976) 405–435.
- [7] R. Gillan, M.D. Cole, A. Linacre, J.W. Thorpe, N.D. Watson, Comparison of *Cannabis sativa* by random amplification of polymorphic DNA (RAPD) and HPLC of cannabinoids; a preliminary study, *Sci. Justice* 35 (3) (1995) 169–177.
- 15 [8] G. Siniscalco Gigliano, P. Caputo, S. Cozzolino, Ribosomal DNA analysis as a tool to identify specimens of *Cannabis sativa* L. of forensic interest, *Sci. Justice* 37 (3) (1997) 171–174.
- [9] G. Siniscalco Gigliano, Identification of *Cannabis sativa* L. (*Cannabaceae*) using restriction profiles of the Internal Transcribed Spacer II (ITS2), *Sci. Justice* 38 (4) (1998) 225–230.
- 20 [10] G. Siniscalco Gigliano, Preliminary data on the usefulness of Internal Transcribed Spacer I (ITS1) sequence in *Cannabis sativa* L. identification, *J. Forensic Sci.* 44 (3) (1999) 475–477.
- [11] M. Kohjyouma, I.J. Lee, O. Iida, K. Kurihara, K. Yamada, Y. Makino, S. Sekita, M. Satake, Intraspecific variation in *Cannabis sativa* L. based on intergenic spacer region of chloroplast DNA, *Biol. Pharm. Bull.* 23 (6) (2000) 727–730.
- 25 [12] S. Gilmore, R. Peakall, J. Robertson, Short tandem repeat (STR) DNA markers are hypervariable and informative in *Cannabis sativa*: implications for forensic investigations, *Forensic Sci. Int.* 131 (2003) 65–74.

- [13] H.M. Hsieh, R.J. Hou, L.C. Tsai, C.S. Wei, S.W. Liu, L.H. Huang, Y.C. Kuo, A. Linacre, J.C.I. Lee, A highly polymorphic STR locus in *Cannabis sativa*, Forensic Sci. Int. 131 (2003) 53–58.
- [14] D. Balding, Forensic applications of microsatellite markers, in D.B. Goldstein and C. Schlötterer (eds.), Microsatellites Evolution and Applications, (1999) Oxford University Press, New York, pp. 198-210.
- [15] K. Weising, H. Nybom, K. Wolff, M. Wieland, DNA fingerprinting in plants and fungi, Boca Ration FL: CRC Press (1995) 51–53.
- [16] S. H. Li, H. Yi-Jiun, J. L. Brown, Isolation of tetranucleotide microsatellites from the Mexican jay *Aphelocoma ultramarina*, Mol. Ecol. 6 (1997) 499–501.
- [17] T. Pearson, Polygyny and extra-pair paternity in a population of southwestern willoflycatchers (*Empidonax traillii extimus*), Northern Arizona University, M.S. thesis (2002)
- [18] L. Cardle, L. Ramsay, D. Milbourne, M. Macaulay, D. Marshall, R. Waugh, Computational and experimental characterization of physically clustered simple sequence repeats in plants, Genetics 156 (2000) 847–854.
- [19] S.D.E. Park, The Excel Microsatellite Toolkit; Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection, University of Dublin, Ph.D. thesis (2001).
- [20] Anonymous, Microsat, the microsatellite distance program, Human population genetics laboratory, Department of genetics, Stanford, CA, USA, (2001).
- [21] J. Felsenstein, PHYLIP (Phylogeny Inference Package) version 3.5c, Seattle, WA, USA, (1993).
- [22] L. Excoffier, P.E. Smouse, J.M. Quattro, Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, Genetics 131 (1992) 179–191.
- [23] R. Peakall, P.E. Smouse, GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia (2001) www.anu.edu.au/BoZo/GenAlEx/

- [24] M. Perez-Reyes, W.R. White, S.A. McDonald, R.E. Hicks, A.R. Jeffcoat, and C.E. Cook, The pharmacological effects of daily marijuana smoking in humans, *Pharmacol. Biochem. Behav.* 40 (1991) 691-694.
- 5 [25] C. Kimpton, D. Fisher, S. Watson, M. Adams, A. Urquhart, J. Lygo, P. Gill. Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci, *Int. J. Legal Med.* 106 (1994) 302-311.
- [26] D. Watts, Genotyping STR loci using an automated DNA sequencer: in P.J. Lincoln, J. Thompson (Eds.), *Forensic DNA Profiling Protocols*, Humana Press, Totowa, New Jersey, 1998, pp. 193-208.
- 10 [27] H. Miller Coyle, A.B. Harper, R.A. Bever, M.A. Sanger, J. Germano-Presby, K.N. Zinnamon, M.A. Menotti-Raymond, V.A. David, Non-human typing: methods and casework applications, *Proceedings of the American Academy of Forensic Sciences 55th Annual Meeting*, (2003) Feb. 17-22, Chicago, Illinois, USA. p. 20.
- 15 [28] H. Miller Coyle, C. Ladd, T. Palmbach, H.C. Lee, The green revolution: botanical contributions to forensics and drug enforcement, *Croatian Med. J.* 42 (3) (2001) 340-345.